

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2004, 09:55:09 ; Search time 565.265 Seconds  
(without alignments)  
10351.438 Million cell updates/sec

Title: US-10-025-145A-3\_COPY\_1560\_1694  
Perfect score: 135  
Sequence: 1 cttgattacagggtgatac.....tcgatcatatcaagcccatg 135

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vl.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	135	100.0	533	6	AR222113	Sequence
2	135	100.0	2018	6	BD227661	BD227661 Synthases
3	135	100.0	2018	6	AR222097	Sequence
4	135	100.0	2018	6	AR240696	Sequence
5	135	100.0	2018	6	AR266971	Sequence
6	135	100.0	2018	6	AR316320	Sequence
7	135	100.0	2018	6	AR338463	Sequence
8	135	100.0	2018	6	AR429869	Sequence
9	135	100.0	2018	8	AGU87909	Sequence
10	119	88.1	1958	8	AY237645	U87909 Abies grand
11	115.8	85.8	2082	8	AF543527	AY237645 Picea sit
12	114.2	84.6	2013	6	AR222136	AF543527 Pinus tae
13	114.2	84.6	2013	8	AGU87910	U87910 Abies grand
14	112.6	83.4	2089	6	BD227677	BD227677 Synthases
15	112.6	83.4	2089	6	AR222098	Sequence
16	112.6	83.4	2089	6	AR240712	Sequence
17	112.6	83.4	2089	6	AR266987	Sequence
18	112.6	83.4	2089	6	AR316336	Sequence
19	112.6	83.4	2089	6	AR338479	Sequence
20	112.6	83.4	2089	6	AR429885	Sequence
21	112.6	83.4	2089	8	AF006193	Sequence
22	111	82.2	2162	8	AF543529	AF006193 Abies gra
23	111	82.2	2429	6	AR222138	AF543529 Pinus tae
24	111	82.2	2429	8	AF139207	AR222138 Sequence
25	109.4	81.0	2198	8	AF369918	AF139207 Abies gra
26	106.2	78.7	1513	6	BD272958	AF369918 Picea abi
27	106.2	78.7	1634	6	BD273051	BD272958 Materials
28	106.2	78.7	2186	6	AR461460	BD273051 Materials
29	104.6	77.5	2196	6	BD227666	AR461460 Picea abi
30	104.6	77.5	2196	6	AR222096	BD227666 Synthases
31	104.6	77.5	2196	6	AR240701	AR222096 Sequence
32	104.6	77.5	2196	6	AR266976	AR240701 Sequence
33	104.6	77.5	2196	6	AR316325	AR266976 Sequence
34	104.6	77.5	2196	6	AR338468	AR316325 Sequence
35	104.6	77.5	2196	6	AR429874	AR338468 Sequence
36	104.6	77.5	2196	8	AGU87908	AR429874 Sequence
37	104.6	77.5	2205	6	AR222116	U87908 Abies grand
38	104	77.0	3306	8	AF326517	AR222116 Sequence
39	103	76.3	696	6	AR222139	AF326517 Abies gra
40	103	76.3	1890	6	AR222146	AR222139 Sequence
41	103	76.3	1893	8	AF139206	AR222146 Sequence
42	102.6	76.0	2150	8	AF369919	AF139206 Abies gra
43	99.8	73.9	2186	6	AR222137	AF369919 Picea abi
44	99.8	73.9	2186	8	AF139205	AR222137 Sequence
45	96.6	71.6	1173	6	BD273050	AF139205 Abies gra
						BD273050 Materials

ALIGNMENTS

RESULT 1	AR222113	Sequence 24	533 bp	DNA	linear	PAT 26-SEP-2002
LOCUS	AR222113	Sequence 24	from patent US 6429014.			
DEFINITION	AR222113	Sequence 24	from patent US 6429014.			
ACCESSION	AR222113	Sequence 24	from patent US 6429014.			
VERSION	AR222113.1	GI:23329487				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						

```
source
1..533
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 135; DB 6; Length 533;
Best Local Similarity 100.0%; Pred. No. 1.1e-33; Mismatches 0; Indels 0; Gaps 0;
Matches 135; Conservative 0;

QY 1 CTTTCGATTACGAGGTGATACGGGTGCTACAAAGCGGACAGGGCTCGTGAGAGAAAGCT 60
Dbb 369 CTTTCGATTACGAGGTGATACGGGTGCTACAAAGCGGACAGGGCTCGTGAGAGAAAGCT 428
QY 61 TCCTCTATATCATGTTATATGAAGACAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 120
Dbb 429 TCCTCTATATCATGTTATATGAAGACAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 488
QY 121 CATATCAACGCCATG 135
Dbb 489 CATATCAACGCCATG 503

RESULT 2
BD227661 2018 bp DNA linear PAT 17-JUL-2003
LOCUS Syntheses.
DEFINITION BD227661
ACCESSION BD227661.1 GI:33037431
VERSION JP 2002526066-A/13.
KEYWORDS Abies grandis
SOURCE Abies grandis
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
REFERENCE 1 (bases 1 to 2018)
AUTHORS Chappell, J., Manna, K.R., Noel, J.P. and Starks, C.M.
TITLE Patent: JP 2002526066-A 13 20-AUG-2002;
JOURNAL UNIVERSITY OF KENTUCKY RESEARCH DEPARTMENT, THE SALK INSTITUTE FOR
BIOLOGICAL STUDIES SECRETARY DEPARTMENT OF HEALTH AND HUMAN
SERVICES
COMMENT OS Abies grandis (giant fir)
PN JP 2002526066-A/13
PD 20-AUG-2002
PF 17-SEP-1999 JP 2000574228
PR 18-SEP-1998 US 60/100993, 22-APR-1999 US 60/130628 PR
23-AUG-1999 US 60/150262
PI JOSEPH CHAPPELL, KATHLEEN R MANNA, JOSEPH P NOEL, COURTNEY M PI
STARKS
PC C12N15/09, A01H5/00, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N5/
PC 10, C12N9/12,
PC C12Q1/25, G01N33/68// (C12N9/12, C12R1:91), (C12N9/12, C12R1:19),
PC C12N15/00,
PC C12N5/00, C12N5/00
PC pinene synthase
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
FEATURES
source
1..2018
/organism="Abies grandis"
/mol_type="genomic DNA"
/db_xref="taxon:46611"

ORIGIN
Query Match 100.0%; Score 135; DB 6; Length 2018;
Best Local Similarity 100.0%; Pred. No. 1.3e-33; Mismatches 0; Indels 0; Gaps 0;
Matches 135; Conservative 0;

QY 1 CTTTCGATTACGAGGTGATACGGGTGCTACAAAGCGGACAGGGCTCGTGAGAGAAAGCT 60
Dbb 1560 CTTTCGATTACGAGGTGATACGGGTGCTACAAAGCGGACAGGGCTCGTGAGAGAAAGCT 1619
QY 61 TCCTCTATATCATGTTATATGAAGACAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 120
Dbb 61 TCCTCTATATCATGTTATATGAAGACAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 120

source
1..533
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 135; DB 6; Length 533;
Best Local Similarity 100.0%; Pred. No. 1.1e-33; Mismatches 0; Indels 0; Gaps 0;
Matches 135; Conservative 0;

QY 1 CTTTCGATTACGAGGTGATACGGGTGCTACAAAGCGGACAGGGCTCGTGAGAGAAAGCT 60
Dbb 369 CTTTCGATTACGAGGTGATACGGGTGCTACAAAGCGGACAGGGCTCGTGAGAGAAAGCT 428
QY 61 TCCTCTATATCATGTTATATGAAGACAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 120
Dbb 429 TCCTCTATATCATGTTATATGAAGACAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 488
QY 121 CATATCAACGCCATG 135
Dbb 489 CATATCAACGCCATG 503

RESULT 3
AR222097 2018 bp DNA linear PAT 26-SEP-2002
LOCUS Sequence 3 from patent US 6429014.
DEFINITION AR222097
ACCESSION AR222097
VERSION AR222097.1 GI:23329471
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2018)
AUTHORS Steele, C.L., Bohlmann, J. and Croteau, R.B.
TITLE Monoterpene synthases from grand fir (Abies grandis)
JOURNAL Patent: US 6429014-A 3 06-AUG-2002;
FEATURES
source
1..2018
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 135; DB 6; Length 2018;
Best Local Similarity 100.0%; Pred. No. 1.3e-33; Mismatches 0; Indels 0; Gaps 0;
Matches 135; Conservative 0;

QY 1 CTTTCGATTACGAGGTGATACGGGTGCTACAAAGCGGACAGGGCTCGTGAGAGAAAGCT 60
Dbb 1560 CTTTCGATTACGAGGTGATACGGGTGCTACAAAGCGGACAGGGCTCGTGAGAGAAAGCT 1619
QY 61 TCCTCTATATCATGTTATATGAAGACAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 120
Dbb 61 TCCTCTATATCATGTTATATGAAGACAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 1679

source
1..533
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 135; DB 6; Length 533;
Best Local Similarity 100.0%; Pred. No. 1.3e-33; Mismatches 0; Indels 0; Gaps 0;
Matches 135; Conservative 0;

QY 1 CTTTCGATTACGAGGTGATACGGGTGCTACAAAGCGGACAGGGCTCGTGAGAGAAAGCT 60
Dbb 1560 CTTTCGATTACGAGGTGATACGGGTGCTACAAAGCGGACAGGGCTCGTGAGAGAAAGCT 1619
QY 61 TCCTCTATATCATGTTATATGAAGACAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 120
Dbb 61 TCCTCTATATCATGTTATATGAAGACAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 120
```

```
|||||
Db 1620 TCCTCTATATCATGTTATATGAAGACAAATCCTGGAGTATCAGAGAAAGATGCTCTCGAT 1679
QY 121 CATATCAACGCCATG 135
Db 1680 CATATCAACGCCATG 1694

RESULT 5
LOCUS AR266971 2018 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 19 from patent US 6495354.
ACCESSION AR266971
VERSION AR266971.1 GI:29696426
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2018)
AUTHORS Chappell,J., Manna,K.R., Noel,J.P. and Starks,C.M.
TITLE Syntheses
JOURNAL Patent: US 6495354-A 19 17-DEC-2002;
FEATURES Location/Qualifiers
source 1..2018
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 135; DB 6; Length 2018;
Best Local Similarity 100.0%; Pred. No. 1.3e-33;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTCGATTACGAGGTGATACCGGTGCTTACAAGCGGACAGGGCTCGTGGAGAAAGACT 60
Db 1560 CTTCGATTACGAGGTGATACCGGTGCTTACAAGCGGACAGGGCTCGTGGAGAAAGACT 1619
QY 61 TCCTCTATATCATGTTATATGAAGACAAATCCTGGAGTATCAGAGAAAGATGCTCTCGAT 120
Db 1620 TCCTCTATATCATGTTATATGAAGACAAATCCTGGAGTATCAGAGAAAGATGCTCTCGAT 1679
QY 121 CATATCAACGCCATG 135
Db 1680 CATATCAACGCCATG 1694

RESULT 6
LOCUS AR316320 2018 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 19 from patent US 6559297.
ACCESSION AR316320
VERSION AR316320.1 GI:31711055
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2018)
AUTHORS Chappell,J., Manna,K.R., Noel,J.P. and Starks,C.M.
TITLE Syntheses
JOURNAL Patent: US 6559297-A 19 06-MAY-2003;
FEATURES Location/Qualifiers
source 1..2018
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 135; DB 6; Length 2018;
Best Local Similarity 100.0%; Pred. No. 1.3e-33;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTCGATTACGAGGTGATACCGGTGCTTACAAGCGGACAGGGCTCGTGGAGAAAGACT 60
Db 1560 CTTCGATTACGAGGTGATACCGGTGCTTACAAGCGGACAGGGCTCGTGGAGAAAGACT 1619
```

```
QY 61 TCCTCTATATCATGTTATATGAAGACAAATCCTGGAGTATCAGAGAAAGATGCTCTCGAT 120
Db 1620 TCCTCTATATCATGTTATATGAAGACAAATCCTGGAGTATCAGAGAAAGATGCTCTCGAT 1679
QY 121 CATATCAACGCCATG 135
Db 1680 CATATCAACGCCATG 1694

RESULT 7
LOCUS AR338463 2018 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 19 from patent US 6569656.
ACCESSION AR338463
VERSION AR338463.1 GI:33725240
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2018)
AUTHORS Chappell,J., Manna,K.R., Noel,J.P. and Starks,C.M.
TITLE Syntheses
JOURNAL Patent: US 6569656-A 19 27-MAY-2003;
FEATURES Location/Qualifiers
source 1..2018
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 135; DB 6; Length 2018;
Best Local Similarity 100.0%; Pred. No. 1.3e-33;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTCGATTACGAGGTGATACCGGTGCTTACAAGCGGACAGGGCTCGTGGAGAAAGACT 60
Db 1560 CTTCGATTACGAGGTGATACCGGTGCTTACAAGCGGACAGGGCTCGTGGAGAAAGACT 1619
QY 61 TCCTCTATATCATGTTATATGAAGACAAATCCTGGAGTATCAGAGAAAGATGCTCTCGAT 120
Db 1620 TCCTCTATATCATGTTATATGAAGACAAATCCTGGAGTATCAGAGAAAGATGCTCTCGAT 1679
QY 121 CATATCAACGCCATG 135
Db 1680 CATATCAACGCCATG 1694

RESULT 8
LOCUS AR429869 2018 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 19 from patent US 6645762.
ACCESSION AR429869
VERSION AR429869.1 GI:40190267
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2018)
AUTHORS Chappell,J., Manna,K.R., Noel,J.P. and Starks,C.M.
TITLE Syntheses
JOURNAL Patent: US 6645762-A 19 11-NOV-2003;
FEATURES Location/Qualifiers
source 1..2018
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 135; DB 6; Length 2018;
Best Local Similarity 100.0%; Pred. No. 1.3e-33;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTCGATTACGAGGTGATACCGGTGCTTACAAGCGGACAGGGCTCGTGGAGAAAGACT 60
Db 1560 CTTCGATTACGAGGTGATACCGGTGCTTACAAGCGGACAGGGCTCGTGGAGAAAGACT 1619
```

QY 61 TCCTCTATATCATGTTATATGAAGACAAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 120  
 Db 1620 TCCTCTATATCATGTTATATGAAGACAAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 1679  
 QY 121 CATATCAAGCCCATG 135  
 Db 1680 CATATCAAGCCCATG 1694

RESULT 9  
 AGU87909  
 LOCUS 2018 bp mRNA linear PLN 24-SEP-1997  
 DEFINITION Abies grandis pinene synthase (AG3.18) mRNA, complete cds.  
 ACCESSION U87909  
 VERSION U87909.1 GI:2411482  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Abies grandis  
 Abies grandis  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.  
 1 (bases 1 to 2018)  
 Bohlmann, J., Steele, C.L. and Croteau, R.  
 Monoterpene synthases from grand fir (Abies grandis). cDNA  
 isolation, characterization, and functional expression of myrcene  
 synthase, (-)-(4S)-limonene synthase, and (-)-(1S,5S)-pinene  
 synthase  
 J. Biol. Chem. 272 (35), 21784-21792 (1997)  
 97413772  
 9268308  
 REFERENCE 2 (bases 1 to 2018)  
 Bohlmann, J., Steele, C.L. and Croteau, R.  
 Direct Submission  
 TITLE  
 JOURNAL  
 Submitted (31-JAN-1997) Institute of Biological Chemistry,  
 Washington State University, Clark Hall, Pullman, WA 99164-6340,  
 USA

FEATURES  
 source  
 1. .2018  
 /organism="Abies grandis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:46611"  
 1. .2018  
 /gene="AG3.18"  
 6. .1892  
 /gene="AG3.18"  
 /note="monoterpene synthase; terpene synthase; monoterpene  
 synthase; terpene cyclase"  
 /codon\_start=1  
 /product="pinene synthase"  
 /protein\_id="AAB71085.1"  
 /db\_xref="GI:2411483"  
 /translation="MALVSTAPLASKSLKSLISSTHKLALSRITPILGMSRRGKS  
 ITFISMSSTVTVDGVRMRMGDFHNLWDDVIQSLPTAYEEKSYLPERAEKLGIV  
 KNFMSMSLEDGELMSPNDLIQRLMTVDLSERLGHRRHFKOSIKKALDYVSYWGEN  
 GICGSRVVDLNLSTALGRLTLRLHGVPSVDFKAFQNGQFSCSENIQDEEIR  
 GVLNLFRAVLIAPGKIMDEARITFTKYLKALQKIPVSSLGREIGDVLGYQWHTYL  
 PREARNYIOVFQDQNTKSYVKSLLKLEAFNIQSLOKRELESILVRWKESG  
 PFEWTCRHRHVEYTLASCIAPFQHSGRFLGFAKTCHLITVLDDMYDTFTGTVDELE  
 LFTATMKRWDPSIDCLPEYMKGVYIAVTVMEMAREAEAGQRTLTLYAREWEAY  
 IDSYMFOEARITGLPSPDEYFENGKVSCHGRISALQPIITMDIPEPDKHLEVDPP  
 SKNDLACALIRLGRDTCYKADPARGEEASSICYNKMPGVSDEADLHINAMISD  
 VIKGLNELLKPDINVPISAKHAFDIARAFHYGYKYRDGYSVANVETKSLVTRTLE  
 SVPL"

ORIGIN  
 Query Match 100.0%; Score 135; DB 8; Length 2018;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-33;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTTCGATTACGAGGTGATACGCGGTGCTACAAGCGGACAGAGGCTCGTGAGAGAAGCT 60  
 Db 1560 CTTCGATTACGAGGTGATACGCGGTGCTACAAGCGGACAGAGGCTCGTGAGAGAAGCT 1619

QY 61 TCCTCTATATCATGTTATATGAAGACAAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 120  
 Db 1620 TCCTCTATATCATGTTATATGAAGACAAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 1679  
 QY 121 CATATCAAGCCCATG 135  
 Db 1680 CATATCAAGCCCATG 1694

RESULT 10  
 AY237645  
 LOCUS 1958 bp mRNA linear PLN 13-SEP-2003  
 DEFINITION Picea sitchensis pinene synthase mRNA, complete cds.  
 ACCESSION AY237645  
 VERSION AY237645.1 GI:34582666  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Picea sitchensis  
 Picea sitchensis  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.  
 1 (bases 1 to 1958)  
 Byun McKay, S.A., Hunter, W.L., Godard, K.A., Wang, S.X., Martin, D.M.,  
 Bohlmann, J. and Plant, A.L.  
 Insect Attack and Wounding Induce Traumatic Resin Duct Development  
 and Gene Expression of (-)-Pinene Synthase in Sitka Spruce  
 Plant Physiol. 133 (1), 368-378 (2003)  
 12970502  
 REFERENCE 2 (bases 1 to 1958)  
 Byun McKay, S.A., Hunter, W.L., Goddard, K.A., Wang, S.X., Martin, D.,  
 Bohlmann, J. and Plant, A.L.  
 Direct Submission  
 TITLE  
 JOURNAL  
 Submitted (17-FEB-2003) Biology, Simon Fraser University, 8888  
 University Drive, Burnaby, BC V5A 1S6, Canada  
 FEATURES  
 Location/Qualifiers  
 1. .1958  
 /organism="Picea sitchensis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3332"  
 /clone="19-2-1/12-3"  
 75. .1958  
 /note="terpene synthase; monoterpene synthase; terpene  
 cyclase; (-)-pinene synthase"  
 /codon\_start=1  
 /product="pinene synthase"  
 /protein\_id="AAP72020.1"  
 /db\_xref="GI:34582667"  
 /translation="MALVSVAPMASRSLHSLSSAHELKTICTIPTILGMSRRGKS  
 ATPMSMSLTITVSDGVQVRMGDFHNLWDDFIQSLSTSYGEPYRERAEKLGIV  
 KNFMSMSLEDGELISPHNDLIQRLVWVDSVRELGIERHKNEIKSALDYVSYWSEK  
 GICGSRVVDLNLSTALGRLTLRLHGVAVSADVLNFKDQNGQFACSPQTEEIRS  
 VLNLYRASLIAPFGEKMBEAEIFSAKYLEESLQKISVLSLOEIRDLVEYGHWTYLP  
 RMEARNHIDVFQDQNTKSCINTEKLELLEAFNIHSLQKRELEYLVRWKDSGS  
 PMAWTCRHRHVEYTLASCIAPFQHSGRFLGFAKCHITITLDDMYDTFTGTVDELEL  
 FTAAMKRWDPSSADCLPEYMKGVYILYDTVNETSREAEKAQGRDITDVARAWDDYL  
 DSNVQAKWIAATGYLTPFAEYFENGKVSCHGRISALQPIITMDIPEPDKHLEVDPP  
 KNDLALASALIRLGRDTCYKADPARGEEASSICYNKMPGVSDEADLHINAMISDV  
 IRLNELLNPNSSVPISSKKHVFDISRAPIHYGYKYRDGYSVANIEKSLVTRTVIDP  
 VTL"

ORIGIN  
 Query Match 88.1%; Score 119; DB 8; Length 1958;  
 Best Local Similarity 92.6%; Pred. No. 2.5e-28;  
 Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 CTTCGATTACGAGGTGATACGCGGTGCTACAAGCGGACAGAGGCTCGTGAGAGAAGCT 60  
 Db 1626 CTTCGACTACGAGGTGATACGCGGTGCTACAAGCGGACAGAGGCTCGTGAGAGAAGCT 1685  
 QY 61 TCCTCTATATCATGTTATATGAAGACAAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 120  
 Db 1686 TCTAGTATATCTTGTATATGAAGACAAATCTCGAGCAACAGAGGAAGATGCTCTCGAT 1745  
 QY 121 CATATCAAGCCCATG 135

```

|||||
Db 1746 CATATCAAGCCATG 1760

RESULT 11
AF543527
LOCUS 2082 bp mRNA linear PIN 10-MAR-2003
DEFINITION Pinus taeda (-)-alpha-pinene synthase mRNA, complete cds.
ACCESSION AF543527
VERSION AF543527.1 GI:28894481
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
cDNA isolation, functional expression, and characterization of
(+)-alpha-pinene synthase and (-)-alpha-pinene synthase from
loblolly pine (Pinus taeda): Stereocontrol in pinene biosynthesis
Arch. Biochem. Biophys. 411 (2), 267-276 (2003)
JOURNAL MEDLINE
PUBMED 12623076
REFERENCE
AUTHORS
TITLE
2 (bases 1 to 2082)
Phillips,M.A., Wildung,M.R., Williams,D.C., Hyatt,D.C. and
Croteau,R.
cDNA isolation, functional expression, and characterization of
(+)-alpha-pinene synthase and (-)-alpha-pinene synthase from
loblolly pine (Pinus taeda): Stereocontrol in pinene biosynthesis
Arch. Biochem. Biophys. 411 (2), 267-276 (2003)
JOURNAL MEDLINE
PUBMED 12623076
REFERENCE
AUTHORS
TITLE
2 (bases 1 to 2082)
Phillips,M.A., Wildung,M.R., Williams,D.C. and Croteau,R.B.
Direct Submission
Submitted (09-SEP-2002) Institute of Biological Chemistry,
Washington State University, PO Box 646340, Pullman, WA 99164-6340,
USA
FEATURES
source Location/Qualifiers
1..2082
/organism="Pinus taeda"
/mol_type="mRNA"
/db_xref="taxon:3352"
32..1921
/note="monoterpene synthase"
/codon_start=1
/product="(-)-alpha-pinene synthase"
/protein_id="AAO61225.1"
/db_xref="GI:28894482"
/translation="MSPVSVISLPSDLCLPSTFIDRSGRELIPLHITIPVAMRROGK
LMTASMSNLRVAVSDDAVIRRGDSFNSLWDDLLQSLSSPYGEPSEYRERAEILIG
EKGKSNMSNDESITPLDLLQRLMWVDSVERLGRHFRKEIKSALDHVYRYS
RTVLNLRASLIAPFGKWDPAESFSAKYLKEALQKIPVSSFSREIGDVLVEWHY
LPRLEARNYIDVFQDTEKNSYMKTEKLEKLEFNIFALQKRELYLVRWVWKS
GSPQWTFORHRRHVEYVTLASCIAPFQHSQPRIGFPAKACHIIITVLDMDYDTFGTDEL
ELFTSAIKRMDPSATECLPEYMKGVYMLVYTNVMSQERADKAQGRITLYCROWEE
YIDAYMEQKWIASGEVPTPEEYENKVSCHRVSAQLPILITDIEPHEVLKEVDI
PSQLNDLASALLRUGDRTRYQADRGEEASCSYMKNPQGTTEEDALNHLNMTIS
DVIGKLWELLKPNSSVPISAKHAFDISRAFHCGYKRYRDGYSVANIEKSLVKRTVI
DPVTI"
ORIGIN
Query Match 85.8%; Score 115.8; DB 8; Length 2082;
Best Local Similarity 91.1%; Pred. No. 2.9e-27;
Matches 123; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 CTTGATTACGAGGTGATACGCGTGTCTACAAAGCGGACAGCGTCTGTGGAGAAGAGCT 60
Db 1589 CTTGATTACGAGGGATACGCGTGTCTACACAGCGGACAGGGCCCGTGGAGAAGAGCT 1648

Qy 61 TCCTCTATATCATGTTATGAAGACAAATCCTGGAGTATCAGAGGAAGATGCTCTCGAT 120
Db 1649 TCGGTATATCTTGTATATGAAGACAAATCCTGGACAAACAGAGGAAGATGCTCTCAAT 1708

Qy 121 CATATCAAGCCCATG 135
Db 1709 CATCTCAAGCCCATG 1723

RESULT 12
AF543527
LOCUS 2082 bp mRNA linear PIN 10-MAR-2003
DEFINITION Pinus taeda (-)-alpha-pinene synthase mRNA, complete cds.
ACCESSION AF543527
VERSION AF543527.1 GI:28894481
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
cDNA isolation, functional expression, and characterization of
(+)-alpha-pinene synthase and (-)-alpha-pinene synthase from
loblolly pine (Pinus taeda): Stereocontrol in pinene biosynthesis
Arch. Biochem. Biophys. 411 (2), 267-276 (2003)
JOURNAL MEDLINE
PUBMED 12623076
REFERENCE
AUTHORS
TITLE
2 (bases 1 to 2082)
Phillips,M.A., Wildung,M.R., Williams,D.C., Hyatt,D.C. and
Croteau,R.
cDNA isolation, functional expression, and characterization of
(+)-alpha-pinene synthase and (-)-alpha-pinene synthase from
loblolly pine (Pinus taeda): Stereocontrol in pinene biosynthesis
Arch. Biochem. Biophys. 411 (2), 267-276 (2003)
JOURNAL MEDLINE
PUBMED 12623076
REFERENCE
AUTHORS
TITLE
2 (bases 1 to 2082)
Phillips,M.A., Bohlmann,J. and Croteau,R.B.
Monoterpene synthases from grand fir (Abies grandis)
Patent: US 6429014-A 64 06-AUG-2002;
Location/Qualifiers
1..2013
/source
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 84.6%; Score 114.2; DB 6; Length 2013;
Best Local Similarity 90.4%; Pred. No. 9.7e-27;
Matches 122; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 CTTGATTACGAGGTGATACGCGTGTCTACAAAGCGGACAGCGTCTGTGGAGAAGAGCT 60
Db 1560 CTTGATTACGAGGTGATACACGCGTGTCTACAAAGCGGACAGGGCCCGTGGAGAAGAGCT 1619

Qy 61 TCCTCTATATCATGTTATGAAGACAAATCCTGGAGTATCAGAGGAAGATGCTCTCGAT 120
Db 1620 TCGTCTATATCATGTTATGAAGACAAATCCTGGATTACGGAAGAGATGCTCTGAAT 1679

Qy 121 CATATCAAGCCCATG 135
Db 1680 CATATCAACTTCATG 1694

RESULT 13
AGU87910
LOCUS 2013 bp mRNA linear PIN 10-AUG-2001
DEFINITION Abies grandis (-)-camphene synthase (AG6.5) mRNA, complete cds.
ACCESSION U87910
VERSION U87910.1 GI:2411484
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
cDNA cloning, characterization, and functional expression of four
new monoterpene synthase members of the Ipsid gene family from grand
fir (Abies grandis)
Arch. Biochem. Biophys. 368 (2), 232-243 (1999)
JOURNAL MEDLINE
PUBMED 99373092
REFERENCE
AUTHORS
TITLE
3 (bases 1 to 2013)
Bohlmann,J., Steele,C.L. and Croteau,R.
Direct Submission
Submitted (31-JAN-1997) Institute of Biological Chemistry,
Washington State University, Clark Hall, Pullman, WA 99164-6340,
USA
FEATURES
source Location/Qualifiers
1..2013
/organism="Abies grandis"

```

```

/molecule="mRNA"
/db_xref="taxon:46611"
1..2013
/gene="Ag6.5"
36..1892
/gene="Ag6.5"
/feature="terpene synthase; monoterpene synthase; terpene cyclase"
/codon_start=1
/product="(-)-camphene synthase"
/protein_id="AAB70707.1"
/db_xref="GI:2411485"
/translation="MALLSITPLVSRSCSSSHEIKALRTIPTILGICRPGKSVASHI
NNCLSVASTDSVQRRVGNHNSLMDDDFIQSLISTPYGADYRERADRLIGEVDIM
ENFKSLDGNLLORLHLLVDVERLGDHFKKIKTKALDYVNSVWNEKIGICGREG
VYDNLNLTALGLRLHGYTVSSDLVNFVKNGQFSTANIQLIEGEIRGVNLFRRA
SLVAFPEKVMDEAETFTKYIREALQKIPASSIISLEIRDLVLEYGWHNLPRLERN
YNDVFGQHTKNNAEKLLEAKLEFNIHFSQERELKHVRWMDKSGSPFEMTFCKHR
HYEYVALASCIAPFQPHGFRGLFTKMSHLITVDDMDYDFVGTVDLELFTATIKRW
PSAMECLPYMGVVMYVHTVNMARVAKQGRDITLVYQAMACFDSYMQEAKW
IATGYLPTFEVLENGKVSASHPALQPIITLIDIPDHLKEVDFFPSKLDLICI
LRLGDRTRYKADRRGEASISCYMKDNPGLTEDALNHINFMIRDAIRLWELL
KPDNSVPIITSKKHAFDISRVHHGYRDRDGYSPANVETKSLVMTVIEPVPL"

ORIGIN
Query Match 84.6%; Score 114.2; DB 8; Length 2013;
Best Local Similarity 90.4%; Pred. No. 9.7e-27;
Matches 122; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CTTGATTACGAGGTGATACGCGTCTACAGGCGGACAGGCGTCTGTGAGAGAAGCT 60
Db 1560 CTTGATTACGAGGTGATACGCGTCTACAGGCGGACAGGCGTCTGTGAGAGAAGCT 1619

QY 61 TCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 120
Db 1620 TCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 1679

QY 121 CATATCAAGCCCATG 135
Db 1680 CATATCAAGCCCATG 1694

RESULT 14
BD227677 2089 bp DNA linear PAT 17-JUL-2003
LOCUS Syntheses.
DEFINITION BD227677.1 GI:33037447
ACCESSION JP 2002526066-A/29.
KEYWORDS Abies grandis
SOURCE Abies grandis
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
Chappell, J., Manna, K.R., Noel, J.P. and Starks, C.M.
Patent: JP 2002526066-A 29 20-AUG-2002;
UNIVERSITY OF KENTUCKY RESEARCH DEPARTMENT, THE SALK INSTITUTE FOR
BIOLOGICAL STUDIES SECRETARY DEPARTMENT OF HEALTH AND HUMAN
SERVICES
OS Abies grandis (giant fir)
PN JP 2002526066-A/29
PD 20-AUG-2002
PF 17-SEP-1999 JP 2000574228
PR 18-SEP-1998 US 60/100993, 22-APR-1999 US 60/130628 PR
23-AUG-1999 US 60/150262
PI JOSEPH CHAPPELL, KATHLEEN R MANNA, JOSEPH P NOEL, COURTNEY M PI
STARKS
PC C12N15/09, A01H5/00, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N5/
PC 10, C12N9/12,
PC C1201/25, G01N33/68// (C12N9/12, C12R1:91), (C12N9/12, C12R1:19),
PC C12N15/00
PC C12N5/00, C12N5/00
```

```

CC (-)-4S-limonene synthase
FH Key Location/Qualifiers
FT CDS (73)..(1983).
FEATURES
source
1..2089
/organism="Abies grandis"
/mol_type="genomic DNA"
/db_xref="taxon:46611"

ORIGIN
Query Match 83.4%; Score 112.6; DB 6; Length 2089;
Best Local Similarity 89.6%; Pred. No. 3.3e-26;
Matches 121; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 CTTGATTACGAGGTGATACGCGTCTACAGGCGGACAGGCGTCTGTGAGAGAAGCT 60
Db 1654 CTTGATTACGAGGTGATACGCGTCTACAGGCGGACAGGCGTCTGTGAGAGAAGCT 1713

QY 61 TCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 120
Db 1714 TCAGCTATATCGTGTATATGAAGACCATCTCTGGATCAATAGAGGAAGATGCTCTCAAT 1773

QY 121 CATATCAAGCCCATG 135
Db 1774 CATATCAAGCCCATG 1788

RESULT 15
AR222098 2089 bp DNA linear PAT 26-SEP-2002
LOCUS Sequence 5 from patent US 6429014.
DEFINITION AR222098
ACCESSION AR222098
VERSION AR222098.1 GI:23329472
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2089)
AUTHORS Steele, C.L., Bohlmann, J. and Croteau, R.B.
TITLE Monoterpene synthases from grand fir (Abies grandis)
JOURNAL Patent: US 6429014-A 5 06-AUG-2002;
FEATURES
source
1..2089
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 83.4%; Score 112.6; DB 6; Length 2089;
Best Local Similarity 89.6%; Pred. No. 3.3e-26;
Matches 121; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 CTTGATTACGAGGTGATACGCGTCTACAGGCGGACAGGCGTCTGTGAGAGAAGCT 60
Db 1654 CTTGATTACGAGGTGATACGCGTCTACAGGCGGACAGGCGTCTGTGAGAGAAGCT 1713

QY 61 TCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 120
Db 1714 TCAGCTATATCGTGTATATGAAGACCATCTCTGGATCAATAGAGGAAGATGCTCTCAAT 1773

QY 121 CATATCAAGCCCATG 135
Db 1774 CATATCAAGCCCATG 1788

Search completed: July 8, 2004, 15:05:42
Job time : 567.265 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2004, 09:46:39 ; Search time 61.0265 Seconds  
(without alignments)  
9397.662 Million cell updates/sec

Title: US-10-025-145A-3\_COPY\_1560\_1694  
Perfect score: 135  
Sequence: 1 cttgattacaggtgatac.....togatcatatcaacgcccatg 135

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 29Jan04:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002s:\*
- 7: geneseqn2003as:\*
- 8: geneseqn2003bs:\*
- 9: geneseqn2003cs:\*
- 10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	135	100.0	533	2	AAX08660 Probe for
2	135	100.0	533	4	Aaf73388 Grand fir
3	135	100.0	2018	2	AAX08644 Pinene sy
4	135	100.0	2018	3	AAX38922 Grand fir
5	135	100.0	2018	4	Aaf73372 Grand fir
6	114.2	84.6	2013	4	Aaf73411 Grand fir
7	112.6	83.4	2089	2	AAX08645 Limonene
8	112.6	83.4	2089	3	AAX38938 Grand fir
9	112.6	83.4	2089	4	Aaf73373 Grand fir
10	111	82.2	2429	4	Aaf73413 Grand fir
11	106.2	78.7	1513	3	AAX69551 Pinus rad
12	106.2	78.7	1634	3	AAX69644 Pinus rad
13	104.6	77.5	2196	2	AAX08643 Myrcene s
14	104.6	77.5	2196	3	AAX38927 Grand fir
15	104.6	77.5	2196	4	Aaf73371 Grand fir
16	104.6	77.5	2205	2	AAX08663 Grand fir
17	104.6	77.5	2205	4	Aaf73391 Grand fir
18	103	76.3	696	4	Aaf73414 Grand fir
19	103	76.3	1890	4	Aaf73421 Grand fir
20	99.8	73.9	2186	4	Aaf73412 Grand fir
21	96.6	71.6	1173	3	AAX69643 Pinus rad
22	66	48.9	622	2	AAX87526 Grand fir
23	66	48.9	2424	2	AAX08654 Grand Fir

24	66	48.9	2424	3	AAA38932	Aaa38932 Grand fir
25	66	48.9	2424	4	Aaf73382	Aaf73382 Grand fir
26	66	48.9	2525	2	AAX87530	Aax87530 E-alpha-b
27	66	48.9	2528	2	AAX87529	Aax87529 Grand fir
28	66	48.9	2528	2	AAH87531	Aax87531 E-alpha-b
29	66	48.9	2571	2	AAX87504	Aax87504 Grand fir
30	56.4	41.8	2700	2	AAT97447	Aat97447 Pacific Y
31	56.4	41.8	2700	3	AAA38931	Aaa38931 Yew taxad
32	52.6	39.0	824	3	AAA69540	Aaa69540 Pinus rad
33	52.6	39.0	1865	3	AAA38933	Aaa38933 Grand fir
34	52.6	39.0	1885	2	AAH87532	Aax87532 Delta-sel
35	52.6	39.0	1885	2	AAH87534	Aax87534 Delta-sel
36	52.6	39.0	1888	2	AAH87533	Aax87533 Delta-sel
37	52.6	39.0	1888	2	AAH87505	Aax87505 Grand fir
38	52.6	39.0	1967	2	AAH87513	Aax87513 Grand fir
39	52.6	39.0	1967	2	AAH08655	Aax08655 Grand Fir
40	52.6	39.0	1967	4	AAH73383	Aaf73383 Grand fir
41	51.6	38.2	2861	3	AAA38937	Aaa38937 Grand Fir
42	35.6	26.4	1416	2	AAX08656	Aax08656 Grand Fir
43	35.6	26.4	1416	4	AAH73384	Aaf73384 Grand fir
44	35.6	26.4	1785	2	AAH87536	Aax87536 Gamma-hum
45	35.6	26.4	1785	2	AAH87537	Aax87537 Gamma-hum

## ALIGNMENTS

RESULT 1  
AAX08660  
ID AAX08660 standard; DNA; 533 BP.  
XX  
AC AAX08660;  
XX  
DT 27-SEP-1999 (first entry)  
XX  
DE Probe for identifying monoterpene synthase sequences.

XX  
KW Myrcene synthase; limonene synthase; pinene synthase; flavour;  
KW monoterpene synthase; aroma; defense; plant seed; oil; meal; primer; PCR;  
ss.

XX Abies grandis.

XX WO9902030-A1.

XX 21-JAN-1999.

XX 10-JUL-1998; 98WO-US014528.

XX 11-JUL-1997; 97US-0052249P.

XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX Bohlmann J, Steele CL, Croteau RB;

XX WPI; 1999-120396/10.

XX New isolated gymnosperm monoterpene synthase DNA - obtained from Grand fir (Abies grandis), used to provide plants with modified production of monoterpenes, e.g. myrcene, limonene or pinene.

XX Example 1; Page 10; 121pp; English.

XX Nucleotide sequences encoding myrcene synthase, limonene synthase and pinene synthase from Grand fir may be incorporated into any organism (e.g. intact plant, animal, microbe), or derived cell culture that produces geranyl diphosphate for the production of the aforementioned enzymes or their products. The sequences when expressed in transfected cells may also be used for the production or modification of flavoured aroma properties, improvement of defense capability, and the alteration of other ecological interactions mediated by myrcene, limonene, pinene, or their derivatives. In particular they can be used for the production of plant seeds for the extraction of oil or meal. cDNAs encoding myrcene

CC synthase, pinene synthase and limonene synthase from Grand Fir were  
 CC isolated and sequenced by the following procedure. Based on comparison of  
 CC sequences of different plant synthase enzymes, four conserved sequences  
 CC were identified for construction of four degenerate primers (AA08646-  
 CC 49). One of these primers (AA08648) was an antisense primer and used for  
 CC PCR with the other three sense primers. Only a combination of primers  
 CC AA08648 and AA08649 generated a PCR product which was 110 base pairs in  
 CC length. The amplified product was used in a plasmid construct to  
 CC transform E.coli XL1-Blue cells and the inserts were sequenced. Four  
 CC probes (AA08650-53) were devised from the sequenced insert and were used  
 CC to screen a Grand Fir cDNA library. One of these clones is the myrcene  
 CC synthase sequence given in AA08643. The other clones identified are  
 CC described in AA08654-56. Based on comparison of these clones three new  
 CC conserved sequences were identified for construction of three degenerate  
 CC primers (AA08657-59). A combination of primers AA08657 and AA08659  
 CC generated a PCR product which was approximately 1020 base pairs in  
 CC length. The amplified product was used in a plasmid construct to  
 CC transform E.coli XL1-Blue cells and the inserts were sequenced. The  
 CC sequence was identical in all transformants and used as a probe  
 CC (AA08660) to screen a Grand Fir cDNA library. Clones were subcloned  
 CC using the primers described in AA08664-X08669 and then sequenced using  
 CC the T3 promoter primer (AA08661) and T7 promoter primer (AA08662).  
 CC Clones which afforded weak hybridisation to the probe sequence were  
 CC identical to the myrcene synthase sequence given in AA08643. Clones  
 CC which gave strong hybridisation signals fell into two distinct groups  
 CC represented by the sequences encoding either limonene synthase (AA08644)  
 CC or pinene synthase (AA08645)

XX  
 SQ Sequence 533 BP; 164 A; 116 C; 120 G; 133 T; 0 U; 0 Other;

Query Match 100.0%; Score 135; DB 2; Length 533;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-36;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTCGATTACGAGTGTATATGAAAGACATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 369 CTTTCGATTACGAGTGTATATGAAAGACATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 428

QY 61 TCCTCTATATCATGTTATATGAAAGACATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 429 TCCTCTATATCATGTTATATGAAAGACATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 488

QY 121 CATATCAAGCCCATG 135  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 489 CATATCAAGCCCATG 503

RESULT 2  
 AAF73388  
 ID AAF73388 standard; DNA; 533 BP.  
 XX AAF73388;  
 AC AAF73388;  
 DT 30-APR-2001 (first entry)  
 XX Grand fir monoterpene synthases probe #3 SEQ ID NO: 24.  
 DE Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
 KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase; probe;  
 KW terpinolene synthase; insect resistance; nutrition; PCR primer; ss.  
 XX Abies grandis.  
 OS Abies grandis.  
 XX WO200107565-A2.  
 DN 01-FEB-2001.  
 XX 24-JUL-2000; 2000WO-US020264.  
 XX 26-JUL-1999; 99US-00360545.  
 XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.

PI Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
 XX WPI; 2001-182782/18.  
 XX New nucleic acid encoding monoterpene synthases, for increasing terpene  
 PT synthesis in plants, e.g. for increasing resistance to pests or for  
 PT treatment of cancer.  
 XX Example 1; Page 133; 175pp; English.

XX The present invention provides the protein and coding sequences of  
 CC monoterpene synthases from the grand fir. These include (-)-camphene  
 CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 CC limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 CC and pinene synthase. The sequences can be used to produce transgenic  
 CC plants expressing high levels of the enzymes, resulting in levels which  
 CC are useful in protecting against and treating cancers, and to confer  
 CC insect resistance on plants

XX Sequence 533 BP; 164 A; 116 C; 120 G; 133 T; 0 U; 0 Other;

Query Match 100.0%; Score 135; DB 4; Length 533;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-36;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTCGATTACGAGTGTATATGAAAGACATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 369 CTTTCGATTACGAGTGTATATGAAAGACATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 428

QY 61 TCCTCTATATCATGTTATATGAAAGACATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 429 TCCTCTATATCATGTTATATGAAAGACATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 488

QY 121 CATATCAAGCCCATG 135  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 489 CATATCAAGCCCATG 503

RESULT 3  
 AAX08644  
 ID AAX08644 standard; cDNA; 2018 BP.  
 XX AAX08644;  
 AC AAX08644;  
 DT 27-SEP-1999 (first entry)  
 XX Pinene synthase gene.  
 DE Myrcene synthase; limonene synthase; pinene synthase; flavour; aroma;  
 KW defense; plant seed; oil; meal; ss.  
 XX Abies grandis.  
 OS Abies grandis.  
 XX Key Location/Qualifiers  
 CDS 6..1892  
 FT /\*tag= a  
 FT /product= "Pinene synthase"  
 XX WO9902030-A1.  
 XX 21-JAN-1999.  
 XX 10-JUL-1998; 98WO-US014528.  
 XX 11-JUL-1997; 97US-0052249P.  
 XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.

PI Bohlmann J, Steele CL, Croteau RB;  
 XX WPI; 1999-120396/10.  
 XX P-PSDB; AAW85701.

PT New isolated gymnosperm monoterpene synthase DNA - obtained from Grand  
PT fir (*Abies grandis*), used to provide plants with modified production of  
XX monoterpenes, e.g. myrcene, limonene or pinene.  
XX  
XX Claim 10; Page 74-77; 121pp; English.  
XX  
CC Nucleotide sequences encoding myrcene synthase, limonene synthase and  
CC pinene synthase from Grand fir may be incorporated into any organism  
CC (e.g. intact plant, animal, microbe), or derived cell culture that  
CC produces geranyl diphosphate for the production of the aforementioned  
CC enzymes or their products. The sequences when expressed in transfected  
CC cells may also be used for the production or modification of flavour and  
CC aroma properties, improvement of defense capability, and the alteration  
CC of other ecological interactions mediated by myrcene, limonene, pinene,  
CC or their derivatives. In particular they can be used for the production  
CC of plant seeds for the extraction of oil or meal  
XX  
XX Sequence 2018 BP; 583 A; 431 C; 479 G; 525 T; 0 U; 0 Other;  
XX  
Query Match 100.0%; Score 135; DB 3; Length 2018;  
Best Local Similarity 100.0%; Pred. No. 7.9e-36;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTTCGATTACGAGTGATACCGGTGCTACAGGCGGACAGGGCTCGTGGAGAGAGCT 60  
Db 1560 CTTTCGATTACGAGTGATACCGGTGCTACAGGCGGACAGGGCTCGTGGAGAGAGCT 1619  
QY 61 TCCTCTATATCATGTTATATGAAGACAATCCTGGAGTATCAGAGGAAGATGCTCTCGAT 120  
Db 1620 TCCTCTATATCATGTTATATGAAGACAATCCTGGAGTATCAGAGGAAGATGCTCTCGAT 1679  
QY 121 CATATCAACGCCATG 135  
Db 1680 CATATCAACGCCATG 1694  
XX  
XX RESULT 4  
XX AAA38922  
XX ID AAA38922 standard; DNA; 2018 BP.  
XX AC AAA38922;  
XX AT 25-AUG-2000 (first entry)  
XX DE Grand fir pinene synthase DNA sequence SEQ ID NO:19.  
XX KW Synthese; protein co-ordinate data; active site; modification; terpenoid;  
KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;  
KW isoprenoid; breeding programme; fragrance; flavour; pheromone;  
KW defensive agent; pigment; antitumour; steroid hormone;  
KW signal transduction pathway; bile acid; affinity purification;  
KW photoreceptor; enzymatic synthesis; nutrient supplement;  
KW immunological reagent; ds.  
XX  
XX *Abies grandis*.  
XX WO200017327-A2.  
XX  
XX 30-MAR-2000.  
XX  
XX 17-SEP-1999; 99WO-US021419.  
XX  
XX 18-SEP-1998; 98US-0100993P.  
XX 22-APR-1999; 99US-0130628P.  
XX 23-AUG-1999; 99US-0150262P.  
XX  
XX (KENT ) UNIV KENTUCKY RES DEPT.  
XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
XX  
XX Chappell J, Manna KR, Noel JP, Starks CW;  
XX WPI; 2000-292839/25.  
XX P-PSDB; AAY90837.  
XX  
XX

XX Novel terpene synthase enzymes, useful for producing terpene  
PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known  
PT enzymes by specific amino acid alterations.  
XX  
XX Disclosure; Page 363-366; 450pp; English.  
XX  
CC The present invention describes an isolated terpene synthase (I)  
CC comprising a region with at least 20% identity to region 265-535 of a 548  
CC amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha  
CC -carbon atoms (alphaC) that have interatomic distances, between each  
CC other, within tabulated ranges, have a centre point (within a sphere of  
CC radius 2.3 Angstrom) within tabulated ranges, and have an ordered  
CC arrangement of R groups (defining aa side chains), excluding specific  
CC tabulated arrangements (tables given in the specification). (I), and  
CC related enzymes, are used to produce a wide range of terpenoids (e.g.  
CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,  
CC flavours, pheromones, defensive agents, pigments, antitumour agents,  
CC components of signal transduction pathways, precursors of steroid  
CC hormones and bile acids, as photoreceptors and as co-factor side chains.  
CC Some synthases with little or no catalytic activity (and nucleic acids  
CC encoding them) are used as controls in the analysis of products formed by  
CC enzymatic synthesis; as nutrient supplements; for affinity purification  
CC of isoprenoids; or to develop immunological reagents or nucleic acids for  
CC monitoring expression of terpene synthase or inheritance of the gene in  
CC plant breeding programs. The new synthases may produce novel terpene  
CC products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent  
CC sequences used in the exemplification of the present invention  
XX  
XX Sequence 2018 BP; 583 A; 431 C; 479 G; 525 T; 0 U; 0 Other;  
XX  
Query Match 100.0%; Score 135; DB 3; Length 2018;  
Best Local Similarity 100.0%; Pred. No. 7.9e-36;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTTCGATTACGAGTGATACCGGTGCTACAGGCGGACAGGGCTCGTGGAGAGAGCT 60  
Db 1560 CTTTCGATTACGAGTGATACCGGTGCTACAGGCGGACAGGGCTCGTGGAGAGAGCT 1619  
QY 61 TCCTCTATATCATGTTATATGAAGACAATCCTGGAGTATCAGAGGAAGATGCTCTCGAT 120  
Db 1620 TCCTCTATATCATGTTATATGAAGACAATCCTGGAGTATCAGAGGAAGATGCTCTCGAT 1679  
QY 121 CATATCAACGCCATG 135  
Db 1680 CATATCAACGCCATG 1694  
XX  
XX RESULT 5  
XX AAF73372  
XX ID AAF73372 standard; cDNA; 2018 BP.  
XX AC AAF73372;  
XX AT 30-APR-2001 (first entry)  
XX DE Grand fir (-)-pinene synthase coding sequence SEQ ID NO: 3.  
XX KW Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
KW terpinolene synthase; insect resistance; nutrition; ss.  
XX  
XX *Abies grandis*.  
XX WO200107565-A2.  
XX  
XX 01-FEB-2001.  
XX  
XX 24-JUL-2000; 2000WO-US020264.  
XX  
XX 26-JUL-1999; 99US-00360545.  
XX  
XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
XX

XX Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
PI WPI; 2001-182782/18.  
XX P-PSDB; AAB69371.  
XX  
XX New nucleic acid encoding monoterpene synthases, for increasing terpene  
PT synthesis in plants, e.g. for increasing resistance to pests or for  
PT treatment of cancer.  
XX  
XX Claim 38; Page 108-110; 175pp; English.  
XX  
XX The present invention provides the protein and coding sequences of  
CC monoterpene synthases from the grand fir. These include (-)-camphene  
CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
CC limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
CC and pinene synthase. The sequences can be used to produce transgenic  
CC plants expressing high levels of the enzymes, resulting in levels which  
CC are useful in protecting against and treating cancers, and to confer  
CC insect resistance on plants  
XX  
XX Sequence 2018 BP; 583 A; 431 C; 479 G; 525 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 135; DB 4; Length 2018;  
Best Local Similarity 100.0%; Pred. No. 7.9e-36;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTGATTACGAGGTGATACGCGGTGCTACAAGCGGACAGGGCTCGTGAGAGAAAGCT 60  
DB 1560 CTTGATTACGAGGTGATACGCGGTGCTACAAGCGGACAGGGCTCGTGAGAGAAAGCT 1619  
QY 61 TCCTCTATATCATGTTATATGAAGACAATCCTCGGAGTATCAGAGGAAGATGCTCTCGAT 120  
DB 1620 TCCTCTATATCATGTTATATGAAGACAATCCTCGGAGTATCAGAGGAAGATGCTCTCGAT 1679  
QY 121 CATATCAAGCCCATG 135  
DB 1680 CATATCAAGCCCATG 1694  
RESULT 6  
AAF73411  
ID AAF73411 standard; cDNA; 2013 BP.  
XX  
XX AAF73411;  
XX  
XX 30-APR-2001 (first entry)  
XX  
XX Grand fir monoterpene synthase coding sequence fragment SEQ ID NO: 64.  
XX  
XX Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
XX myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
XX terpinolene synthase; insect resistance; nutrition; ss.  
XX  
XX Abies grandis.  
XX  
XX WO200107565-A2.  
XX  
XX 01-FEB-2001.  
XX  
XX 24-JUL-2000; 2000WO-US020264.  
XX  
XX 26-JUL-1999; 99US-00360545.  
XX  
XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
XX  
XX Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
PI WPI; 2001-182782/18.  
XX P-PSDB; AAB69390.  
XX  
XX New nucleic acid encoding monoterpene synthases, for increasing terpene  
PT synthesis in plants, e.g. for increasing resistance to pests or for

PT treatment of cancer.  
XX  
XX Claim 8; Page 147-149; 175pp; English.  
XX  
XX The present invention provides the protein and coding sequences of  
CC monoterpene synthases from the grand fir. These include (-)-camphene  
CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
CC limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
CC and pinene synthase. The sequences can be used to produce transgenic  
CC plants expressing high levels of the enzymes, resulting in levels which  
CC are useful in protecting against and treating cancers, and to confer  
CC insect resistance on plants  
XX  
XX Sequence 2013 BP; 591 A; 432 C; 464 G; 526 T; 0 U; 0 Other;  
SQ  
Query Match 84.6%; Score 114.2; DB 4; Length 2013;  
Best Local Similarity 90.4%; Pred. No. 1.1e-28;  
Matches 122; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
QY 1 CTTGATTACGAGGTGATACGCGGTGCTACAAGCGGACAGGGCTCGTGAGAGAAAGCT 60  
DB 1560 CTTGATTACGAGGTGATACGCGGTGCTACAAGCGGACAGGGCTCGTGAGAGAAAGCT 1619  
QY 61 TCCTCTATATCATGTTATATGAAGACAATCCTCGGAGTATCAGAGGAAGATGCTCTCGAT 120  
DB 1620 TCCTCTATATCATGTTATATGAAGACAATCCTCGGAGTATCAGAGGAAGATGCTCTCGAT 1679  
QY 121 CATATCAAGCCCATG 135  
DB 1680 CATATCAAGCCCATG 1694  
RESULT 7  
AAX08645  
ID AAX08645 standard; cDNA; 2089 BP.  
XX  
XX AAX08645;  
XX  
XX 27-SEP-1999 (first entry)  
XX  
XX Limonene synthase gene.  
XX  
XX Myrcene synthase; limonene synthase; pinene synthase; flavour; aroma;  
XX defense; plant seed; oil; meal; ss.  
XX  
XX Abies grandis.  
XX  
XX Key Location/Qualifiers  
XX CDS 73..1986  
XX /\*tag= a  
XX /product= "Limonene synthase"  
XX  
XX WO9902030-A1.  
XX  
XX 21-JAN-1999.  
XX  
XX 10-JUL-1998; 98WO-US014528.  
XX  
XX 11-JUL-1997; 97US-0052249P.  
XX  
XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
XX  
XX Bohlmann J, Steele CL, Croteau RB;  
PI WPI; 1999-120396/10.  
XX P-PSDB; AAW85702.  
XX  
XX New isolated gymnosperm monoterpene synthase DNA - obtained from Grand  
PT fir (Abies grandis), used to provide plants with modified production of  
PT monoterpenes, e.g. myrcene, limonene or pinene.  
XX  
XX Claim 11; Page 79-82; 121pp; English.  
XX

CC Nucleotide sequences encoding myrene synthase, limonene synthase and  
 CC pinene synthase from Grand fir may be incorporated into any organism  
 CC (e.g. intact plant, animal, microbe), or derived cell culture that  
 CC produces geranyl diphosphate for the production of the aforementioned  
 CC enzymes or their products. The sequences when expressed in transfected  
 CC cells may also be used for the production or modification of flavour and  
 CC aroma properties, improvement of defense capability, and the alteration  
 CC of other ecological interactions mediated by myrcene, limonene, pinene,  
 CC or their derivatives. In particular they can be used for the production  
 CC of plant seeds for the extraction of oil or meal  
 XX  
 SQ Sequence 2089 BP; 638 A; 415 C; 450 G; 586 T; 0 U; 0 Other;  
 Query Match 83.4%; Score 112.6; DB 2; Length 2089;  
 Best Local Similarity 89.6%; Pred. No. 3.8e-28;  
 Matches 121; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 CTTGATTACGAGGTGATACCGGCTGTACAGGCGGACAGGCTCGTGGAGAGAGCT 60  
 Db 1654 CTTGACTACGAGGTGACACCGCTGTACAGGCGGATAGGCGCGTGGAGAGAGCT 1713  
 QY 61 TCCTCTATATCATGTTATATGAAAGACAAATCCTGAGTATCAGAGGAGATGCTCTCGAT 120  
 Db 1714 TCAGTATATGCTGTATATGAAAGACCAATCCTGGATCAATAGAGAGATGCTCTCAAT 1773  
 QY 121 CATATCAACGCCATG 135  
 Db 1774 CATATCAACGCCATG 1788  
 RESULT 8  
 ID AAA38938 standard; DNA; 2089 BP.  
 AC AAA38938;  
 XX  
 DT 25-AUG-2000 (first entry)  
 DE Grand fir limonene synthase DNA sequence SEQ ID NO:57.  
 KW Synthese; protein co-ordinate data; active site; modification; terpenoid;  
 KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;  
 KW isoprenoid; breeding programme; fragrance; flavour; pheromone;  
 KW defensive agent; pigment; antitumour; steroid hormone;  
 KW signal transduction pathway; bile acid; affinity purification;  
 KW photoreceptor; enzymatic synthesis; nutrient supplement;  
 KW immunological reagent; ds.  
 XX  
 OS Abies grandis.  
 XX  
 XX WO200017327-A2.  
 XX  
 XX 30-MAR-2000.  
 XX  
 PF 17-SEP-1999; 99WO-US021419.  
 XX  
 PR 18-SEP-1998; 98US-0100993P.  
 PR 22-APR-1999; 99US-0130628P.  
 PR 23-AUG-1999; 99US-0150262P.  
 XX  
 XX (KENT) UNIV KENTUCKY RES DEPT.  
 PA (SALK) SALK INST BIOLOGICAL STUDIES.  
 XX  
 XX Chappell J, Manna KR, Noel JP, Starks CM;  
 XX  
 XX WPI; 2000-292839/25.  
 DR P-PSDB; AAY90859.  
 XX  
 XX Novel terpene synthase enzymes, useful for producing terpene  
 PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known  
 PT enzymes by specific amino acid alterations.  
 XX  
 PS Disclosure; Page 445-448; 450pp; English.

XX The present invention describes an isolated terpene synthase (I)  
 CC comprising a region with at least 20% identity to region 265-535 of a 548  
 CC amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha  
 CC -carbon atoms (alphaC) that have interatomic distances, between each  
 CC other, within tabulated ranges, have a centre point (within a sphere of  
 CC radius 2.3 Angstrom) within tabulated ranges, and have an ordered  
 CC arrangement of R groups (defining aa side chains), excluding specific  
 CC tabulated arrangements (tables given in the specification). (I), and  
 CC related enzymes, are used to produce a wide range of terpenoids (e.g.  
 CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,  
 CC flavours, pheromones, defensive agents, pigments, antitumour agents,  
 CC components of signal transduction pathways, precursors of steroid  
 CC hormones and bile acids, as photoreceptors and as co-factor side chains.  
 CC Some synthases with little or no catalytic activity (and nucleic acids  
 CC encoding them) are used as controls in the analysis of products formed by  
 CC enzymatic synthesis; as nutrient supplements; for affinity purification  
 CC of isoprenoids; or to develop immunological reagents or nucleic acids for  
 CC monitoring expression of terpene synthase or inheritance of the gene in  
 CC plant breeding programs. The new synthases may produce novel terpene  
 CC products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 2089 BP; 638 A; 415 C; 450 G; 586 T; 0 U; 0 Other;  
 Query Match 83.4%; Score 112.6; DB 3; Length 2089;  
 Best Local Similarity 89.6%; Pred. No. 3.8e-28;  
 Matches 121; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 CTTGATTACGAGGTGATACCGGCTGTACAGGCGGACAGGCTCGTGGAGAGAGCT 60  
 Db 1654 CTTGACTACGAGGTGACACCGCTGTACAGGCGGATAGGCGCGTGGAGAGAGCT 1713  
 QY 61 TCCTCTATATCATGTTATATGAAAGACAAATCCTGAGTATCAGAGGAGATGCTCTCGAT 120  
 Db 1714 TCAGTATATGCTGTATATGAAAGACCAATCCTGGATCAATAGAGAGATGCTCTCAAT 1773  
 QY 121 CATATCAACGCCATG 135  
 Db 1774 CATATCAACGCCATG 1788  
 RESULT 9  
 ID AAF73373 standard; cDNA; 2089 BP.  
 XX AAF73373;  
 AC AAF73373;  
 XX  
 DT 30-APR-2001 (first entry)  
 DE Grand fir (-)-limonene synthase coding sequence SEQ ID NO: 5.  
 XX  
 XX Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
 KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
 KW terpinolene synthase; insect resistance; nutrition; ss.  
 XX  
 OS Abies grandis.  
 XX  
 XX WO200107565-A2.  
 XX  
 XX 01-FEB-2001.  
 XX  
 XX 24-JUL-2000; 2000WO-US020264.  
 XX  
 XX 26-JUL-1999; 99US-00360545.  
 XX  
 XX (UNIW) UNIV WASHINGTON STATE RES FOUND.  
 XX  
 XX Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
 XX  
 XX WPI; 2001-182782/18.  
 DR P-PSDB; AAB69372.  
 XX



CC therapeutic effects, including direct application in diseased organisms  
 CC or indirect application by transgenic organisms and in fermentation and  
 CC chemical processing industries involving isoprenoids. In plant  
 CC applications, manipulating isoprenoid pathways or isoprenoid composition  
 CC may, for example, affect plant development, pest resistance, and the  
 CC value of extractives (e.g. pinene and myrcene). The ubiquitous and varied  
 CC roles of isoprenoids make the polynucleotides attractive targets for  
 CC biotechnical applications in a variety of fields. AAA69527 to AAA69690  
 CC and AAB18004 to AAB18143 represent Eucalyptus grandis and Pinus radiata  
 CC polynucleotides and proteins used in the exemplification of the present  
 CC invention  
 XX  
 SQ Sequence 1513 BP; 437 A; 316 C; 334 G; 426 T; 0 U; 0 Other;

Query Match 78.7%; Score 106.2; DB 3; Length 1513;  
 Best Local Similarity 86.7%; Pred. No. 5.3e-26;  
 Matches 117; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 CTTGATTACGAGTGATACCGGTGCTACAGCGGACAGGGCTCTGGAGAGAGCT 60  
 Db 1057 CTTGATTACGAGTGATACCGGTGCTACAGCGGACAGGGCTCTGGAGAGAGCT 1116  
 QY 61 TCCTCTATATCATGTTATATGAAGACAACTCTGGAGTATCAGAGGAGATGCTCTCGAT 120  
 Db 1117 TCGTGTATATCTGCTATATGAAGAAATCCCGGTCAACAGGAGATGCTCTCGAT 1176  
 QY 121 CATATCAACGCCATG 135  
 Db 1177 CATATCAACGGTATG 1191

RESULT 12  
 AAA69644  
 ID AAA69644 standard; cDNA; 1634 BP.  
 XX  
 AC AAA69644;  
 XX  
 DT 08-NOV-2000 (first entry)  
 DE  
 DE Pinus radiata pinene synthase cDNA SEQ ID NO:118.  
 XX  
 KW Eucalyptus grandis; Pinus radiata; modification; isoprenoid; plant;  
 KW metabolism; isoprenoid biosynthetic pathway; terpenoid; steroid;  
 KW genome mapping; physical mapping; positional cloning; forestry;  
 KW agriculture; medicine; fermentation; plant development; pest resistance;  
 KW pinene; myrcene; Monterey pine; ss.  
 XX  
 KW Pinus radiata.  
 OS  
 OS WO20036081-A2.  
 EN  
 XX  
 PD 22-JUN-2000.  
 XX  
 XX 16-DEC-1999; 99WO-NZ000219.  
 XX  
 XX 17-DEC-1998; 98US-00215504.  
 PR  
 PR 29-JUL-1999; 99US-0146441P.  
 XX

(GENE-) GENESIS RES & DEV CORP LTD.  
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
 XX  
 XX Havukkala IJ;  
 XX  
 DR WPI; 2000-431575/37.  
 DR P-PSDB; AAB18044.  
 XX

New plant polynucleotides encoding polypeptides involved in the  
 PT production and modification of isoprenoids, useful in forestry and  
 PT agriculture for manipulation of isoprenoid metabolism.  
 XX  
 XX Claim 1; Page 89; 164pp; English.  
 XX  
 XX The present invention describes plant polynucleotides encoding

CC polypeptides involved in the production and modification of isoprenoids,  
 CC such as terpenoid and steroid compounds. The polynucleotides are used in  
 CC genome mapping, in physical mapping and in positional cloning of genes.  
 CC The polynucleotides and polypeptides are useful in forestry and  
 CC agriculture for manipulation of isoprenoid metabolism, in medicine for  
 CC therapeutic effects, including direct application in diseased organisms  
 CC or indirect application by transgenic organisms and in fermentation and  
 CC chemical processing industries involving isoprenoids. In plant  
 CC applications, manipulating isoprenoid pathways or isoprenoid composition  
 CC may, for example, affect plant development, pest resistance, and the  
 CC value of extractives (e.g. pinene and myrcene). The ubiquitous and varied  
 CC roles of isoprenoids make the polynucleotides attractive targets for  
 CC biotechnical applications in a variety of fields. AAA69527 to AAA69690  
 CC and AAB18004 to AAB18143 represent Eucalyptus grandis and Pinus radiata  
 CC polynucleotides and proteins used in the exemplification of the present  
 CC invention  
 XX  
 SQ Sequence 1634 BP; 485 A; 328 C; 355 G; 466 T; 0 U; 0 Other;

Query Match 78.7%; Score 106.2; DB 3; Length 1634;  
 Best Local Similarity 86.7%; Pred. No. 5.4e-26;  
 Matches 117; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 CTTGATTACGAGTGATACCGGTGCTACAGCGGACAGGGCTCTGGAGAGAGCT 60  
 Db 1057 CTTGATTACGAGTGATACCGGTGCTACAGCGGACAGGGCTCTGGAGAGAGCT 1116  
 QY 61 TCCTCTATATCATGTTATATGAAGACAACTCTGGAGTATCAGAGGAGATGCTCTCGAT 120  
 Db 1117 TCGTGTATATCTGCTATATGAAGAAATCCCGGTCAACAGGAGATGCTCTCGAT 1176  
 QY 121 CATATCAACGCCATG 135  
 Db 1177 CATATCAACGGTATG 1191

RESULT 13  
 AAX08643  
 ID AAX08643 standard; cDNA; 2196 BP.  
 XX  
 AC AAX08643;  
 XX  
 DT 27-SEP-1999 (first entry)  
 DE  
 DE Myrcene synthase gene.  
 XX  
 KW Myrcene synthase; limonene synthase; pinene synthase; flavour; aroma;  
 KW defense; plant seed; oil; meal; ss.  
 XX  
 OS Abies grandis.  
 OS  
 OS  
 FH Key Location/Qualifiers  
 FT CDS 69..1952  
 FT /\*tag= a  
 FT /product= "Myrcene synthase"  
 XX  
 XX WO9902030-A1.  
 XX  
 XX 21-JAN-1999.  
 XX  
 XX 10-JUL-1998; 98WO-US014528.  
 XX  
 XX 11-JUL-1997; 97US-0052249P.  
 PR  
 PR (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 XX  
 XX Bohlmann J, Steele CL, Croteau RB;  
 XX WPI; 1999-120396/10.  
 DR  
 DR P-PSDB; AAW85700.  
 XX  
 XX New isolated gymnosperm monoterpane synthase DNA - obtained from Grand  
 PT fir (Abies grandis), used to provide plants with modified production of



```

XX WP1; 2001-182782/18.
DR P-PSDB; AAB69370.
XX
XX PT New nucleic acid encoding monoterpene synthases, for increasing terpene
PT synthesis in plants, e.g. for increasing resistance to pests or for
PT treatment of cancer.
XX
XX PS Claim 28; Page 103-106; 175pp; English.
XX
XX CC The present invention provides the protein and coding sequences of
CC monoterpene synthases from the grand fir. These include (-)-camphene
CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-
CC limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase
CC and pinene synthase. The sequences can be used to produce transgenic
CC plants expressing high levels of the enzymes, resulting in levels which
CC are useful in protecting against and treating cancers, and to confer
CC insect resistance on plants
XX
SQ Sequence 2196 BP; 684 A; 416 C; 468 G; 628 T; 0 U; 0 Other;
Query Match 77.5%; Score 104.6; DB 4; Length 2196;
Best Local Similarity 85.9%; Pred. No. 2.1e-25;
Matches 116; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Qy 1 CTTGATTACGAGGTGATACCGGTGCTACAAAGCGGACAGGGCTCGTGAGAGAAGCT 60
Db 1620 CTTGCGGTACGAGGTGACACACGCTGCTACAAAGCGGATAGGGATCGTGTGAAGAAGCT 1679
Qy 61 TCCTCTATATCATGTTATATGAAGACAACTCTGGAGTATCAGAGGAGATGCTCTCGAT 120
Db 1680 TCGTGTATATCATGTTATATGAAGACAACTCTGGATCAACCGAAGAAGATGCCCTCAAT 1739
Qy 121 CATATCAACGCCATG 135
Db 1740 CATATCAATGCCATG 1754

```

Search completed: July 8, 2004, 12:35:39  
Job time : 63.0265 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2004, 12:09:59 ; Search time 70.7053 Seconds  
(without alignments)  
9305.868 Million cell updates/sec

Title: US-10-025-145A-3\_COPY\_1560\_1694

Perfect score: 135  
Sequence: 1 cttcgattacagggtgatac.....tcgatcatatcaagccatg 135

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3183909 seqs, 2436941669 residues

Total number of hits satisfying chosen parameters: 6367818

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	135	100.0	533	15	US-10-025-145A-24
2	135	100.0	2018	9	US-09-887-586A-19
3	135	100.0	2018	9	US-09-903-012-19
4	135	100.0	2018	10	US-09-900-797-19
5	135	100.0	2018	13	US-09-893-820-19
6	135	100.0	2018	14	US-10-041-007-21
7	135	100.0	2018	15	US-10-025-145A-31
8	114.2	84.6	2013	15	US-10-025-145A-64
9	112.6	83.4	2089	9	US-09-887-586A-57
10	112.6	83.4	2089	9	US-09-903-012-57
11	112.6	83.4	2089	10	US-09-900-797-57
12	112.6	83.4	2089	13	US-09-893-820-57
13	112.6	83.4	2089	14	US-10-041-007-23
14	112.6	83.4	2089	15	US-10-025-145A-5

15	111	82.2	2429	14	US-10-041-007-27	Sequence 27, Appl
16	111	82.2	2429	15	US-10-025-145A-68	Sequence 68, Appl
17	104.6	77.5	2196	9	US-09-887-586A-29	Sequence 29, Appl
18	104.6	77.5	2196	9	US-09-903-012-29	Sequence 29, Appl
19	104.6	77.5	2196	10	US-09-900-797-29	Sequence 29, Appl
20	104.6	77.5	2196	13	US-09-893-820-29	Sequence 29, Appl
21	104.6	77.5	2196	14	US-10-041-007-25	Sequence 25, Appl
22	104.6	77.5	2196	15	US-10-025-145A-1	Sequence 1, Appl
23	104.6	77.5	2205	15	US-10-025-145A-31	Sequence 31, Appl
24	103	76.3	696	15	US-10-025-145A-70	Sequence 70, Appl
25	103	76.3	1890	15	US-10-025-145A-77	Sequence 77, Appl
26	99.8	73.9	2186	15	US-10-025-145A-66	Sequence 66, Appl
27	66	48.9	2424	9	US-09-887-586A-45	Sequence 45, Appl
28	66	48.9	2424	9	US-09-903-012-45	Sequence 45, Appl
29	66	48.9	2424	10	US-09-900-797-45	Sequence 45, Appl
30	66	48.9	2424	13	US-09-893-820-45	Sequence 45, Appl
31	66	48.9	2424	14	US-10-041-007-14	Sequence 14, Appl
32	66	48.9	2424	15	US-10-025-145A-15	Sequence 15, Appl
33	66	48.9	2528	14	US-10-041-007-13	Sequence 13, Appl
34	56.4	41.8	2700	9	US-09-887-586A-43	Sequence 43, Appl
35	56.4	41.8	2700	9	US-09-903-012-43	Sequence 43, Appl
36	56.4	41.8	2700	10	US-09-900-797-43	Sequence 43, Appl
37	56.4	41.8	2700	12	US-10-041-018-363	Sequence 363, Appl
38	56.4	41.8	2700	13	US-09-893-820-43	Sequence 43, Appl
39	52.6	39.0	1865	9	US-09-887-586A-47	Sequence 47, Appl
40	52.6	39.0	1865	9	US-09-903-012-47	Sequence 47, Appl
41	52.6	39.0	1865	10	US-09-900-797-47	Sequence 47, Appl
42	52.6	39.0	1865	13	US-09-893-820-47	Sequence 47, Appl
43	52.6	39.0	1865	14	US-10-041-007-17	Sequence 17, Appl
44	52.6	39.0	1967	15	US-10-025-145A-17	Sequence 17, Appl
45	51.6	38.2	2861	9	US-09-887-586A-55	Sequence 55, Appl

## ALIGNMENTS

### RESULT 1

US-10-025-145A-24  
; Sequence 24, Application US/10025145A  
; Publication No. US20030175861A1  
; GENERAL INFORMATION:  
; APPLICANT: Croteau, Rodney B.  
; APPLICANT: Rohlmann, Joerg  
; APPLICANT: Steele, Christopher L.  
; APPLICANT: Phillips, Michael A.  
; TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)  
; FILE REFERENCE: WSUR118414  
; CURRENT APPLICATION NUMBER: US/10/025,145A  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: US 09/360,545  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: PCT/US98/14528  
; PRIOR FILING DATE: 1998-07-10  
; PRIOR APPLICATION NUMBER: US 60/052,249  
; PRIOR FILING DATE: 1997-07-11  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 533  
; TYPE: DNA  
; ORGANISM: Abies Grandis

Query Match	100.0%	Score 135;	DB 15;	Length 533;
Best Local Similarity	100.0%	Pred No. 4.7e-37;		
Matches 135;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CTTCGATTACAGGTGATACCGGTGCTACAGCGGACAGGCTCGTGAGAGAGCT	60	
Db	369	CTTCGATTACAGGTGATACCGGTGCTACAGCGGACAGGCTCGTGAGAGAGCT	428	
Qy	61	TCCTCTATATCAGTTCATATATGAAGCAATCTCTGGAGTATCAGAGAGATGCTTCGAT	120	

Db 429 TCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 488  
QY 121 CATATCAAGCCATG 135  
Db 489 CATATCAAGCCATG 503

## RESULT 2

US-09-887-586A-19  
; Sequence 19, Application US/09887586A  
; Patent No. US20020094556A1  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. US20020094556A11, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/887,586A  
; PRIOR FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: 09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 19  
; LENGTH: 2018  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (6)...(1889)  
; OTHER INFORMATION: pinene synthase  
US-09-887-586A-19

Query Match 100.0%; Score 135; DB 9; Length 2018;  
Best Local Similarity 100.0%; Pred. No. 8.4e-37;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTTGATTACGAGGTGATACGGGTCTCTACAGCGCGGACAGGGCTCTGTGGAGAAGCT 60  
Db 1560 CTTTGATTACGAGGTGATACGGGTCTCTACAGCGCGGACAGGGCTCTGTGGAGAAGCT 1619  
QY 61 TCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 120  
Db 1620 TCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 1679  
QY 121 CATATCAAGCCATG 135  
Db 1680 CATATCAAGCCATG 1694

## RESULT 3

US-09-903-012-19  
; Sequence 19, Application US/09903012  
; Patent No. US20020094557A1  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. US20020094557A11, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/903,012  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/100,993  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/130,628

; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 19  
; LENGTH: 2018  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (6)...(1889)  
; OTHER INFORMATION: pinene synthase  
US-09-903-012-19

Query Match 100.0%; Score 135; DB 9; Length 2018;  
Best Local Similarity 100.0%; Pred. No. 8.4e-37;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTTGATTACGAGGTGATACGGGTCTCTACAGCGCGGACAGGGCTCTGTGGAGAAGCT 60  
Db 1560 CTTTGATTACGAGGTGATACGGGTCTCTACAGCGCGGACAGGGCTCTGTGGAGAAGCT 1619  
QY 61 TCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 120  
Db 1620 TCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 1679  
QY 121 CATATCAAGCCATG 135  
Db 1680 CATATCAAGCCATG 1694

## RESULT 4

US-09-900-797-19  
; Sequence 19, Application US/09900797  
; Publication No. US20030087406A1  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. US20030087406A11, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/900,797  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: US/09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 19  
; LENGTH: 2018  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (6)...(1889)  
; OTHER INFORMATION: pinene synthase  
US-09-900-797-19

Query Match 100.0%; Score 135; DB 10; Length 2018;  
Best Local Similarity 100.0%; Pred. No. 8.4e-37;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTTGATTACGAGGTGATACGGGTCTCTACAGCGCGGACAGGGCTCTGTGGAGAAGCT 60  
Db 1560 CTTTGATTACGAGGTGATACGGGTCTCTACAGCGCGGACAGGGCTCTGTGGAGAAGCT 1619  
QY 61 TCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 120

```
Db 1620 TCCTCTATATCATGTTATATGAAGACAATCCTGGAGTATCAGAGGAAGATGCTCTCGAT 1679
QY 121 CATATCAACGCCATG 135
Db 1680 CATATCAACGCCATG 1694

RESULT 5
US-09-893-820-19
; Sequence 19, Application US/09893820
; Publication No. US20040053386A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20040053386A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6)...(1889)
; OTHER INFORMATION: pinene synthase
US-09-893-820-19

Query Match 100.0%; Score 135; DB 13; Length 2018;
Best Local Similarity 100.0%; Pred. No. 8.4e-37;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCGATTACGAGGTGATACGCGGTGCTACAGGCGGACAGGGCTCGTGGAGAGAAGCT 60
Db 1560 CTTCGATTACGAGGTGATACGCGGTGCTACAGGCGGACAGGGCTCGTGGAGAGAAGCT 1619
QY 61 TCCTCTATATCATGTTATATGAAGACAATCCTGGAGTATCAGAGGAAGATGCTCTCGAT 120
Db 1620 TCCTCTATATCATGTTATATGAAGACAATCCTGGAGTATCAGAGGAAGATGCTCTCGAT 1679
QY 121 CATATCAACGCCATG 135
Db 1680 CATATCAACGCCATG 1694

RESULT 6
US-10-041-007-21
; Sequence 21, Application US/10041007
; Publication No. US20020164736A1
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seiichi P.T.
; APPLICANT: Schepmann, Hala G.
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase
; FILE REFERENCE: P02081US1
; CURRENT APPLICATION NUMBER: US/10/041,007
; PRIOR FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/259,881
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21

Db 1620 TCCTCTATATCATGTTATATGAAGACAATCCTGGAGTATCAGAGGAAGATGCTCTCGAT 1679
QY 121 CATATCAACGCCATG 135
Db 1680 CATATCAACGCCATG 1694

RESULT 7
US-10-025-145A-3
; Sequence 3, Application US/10025145A
; Publication No. US20030175861A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Bohlmann, Joerg
; APPLICANT: Steele, Christopher L.
; APPLICANT: Phillips, Michael A.
; TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)
; FILE REFERENCE: WSUR118414
; CURRENT APPLICATION NUMBER: US/10/025,145A
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 09/360,545
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: PCT/US98/14528
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: US 60/052,249
; PRIOR FILING DATE: 1997-07-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Abies Grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6)...(1892)
; OTHER INFORMATION:
US-10-025-145A-3

Query Match 100.0%; Score 135; DB 15; Length 2018;
Best Local Similarity 100.0%; Pred. No. 8.4e-37;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCGATTACGAGGTGATACGCGGTGCTACAAAGCGGACAGGGCTCGTGGAGAGAAGCT 60
Db 1560 CTTCGATTACGAGGTGATACGCGGTGCTACAAAGCGGACAGGGCTCGTGGAGAGAAGCT 1619
QY 61 TCCTCTATATCATGTTATATGAAGACAATCCTGGAGTATCAGAGGAAGATGCTCTCGAT 120
Db 1620 TCCTCTATATCATGTTATATGAAGACAATCCTGGAGTATCAGAGGAAGATGCTCTCGAT 1679
QY 121 CATATCAACGCCATG 135
Db 1680 CATATCAACGCCATG 1694

RESULT 8
US-10-025-145A-64
; Sequence 64, Application US/10025145A
```

```
; Publication No. US20030175861A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Bohlmann, Joerg
; APPLICANT: Steele, Christopher L.
; APPLICANT: Phillips, Michael A.
; TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)
; FILE REFERENCE: WSUK118414
; CURRENT APPLICATION NUMBER: US/10/025,145A
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 09/360,545
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: PCT/US98/14528
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: US 60/052,249
; PRIOR FILING DATE: 1997-07-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
; LENGTH: 2013
; TYPE: DNA
; ORGANISM: Abies Grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (36)..(1889)
; OTHER INFORMATION:
; US-10-025-145A-64

Query Match      84.6%; Score 114.2; DB 15; Length 2013;
Best Local Similarity 90.4%; Pred. No. 1.7e-29;
Matches 122; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY      1  CTTGATTACGAGGTGATACGCGGTCTACAAGCGGACAGGGCTCGTGAGAGAAGCT 60
Db      1560  CTTGATTACGAGGTGATACGCGGTCTACAAGCGGACAGGGCTCGTGAGAGAAGCT 1619

QY      61  TCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 120
Db      1620  TCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 1679

QY      121  CATATCAACGCCCATG 135
Db      1680  CATATCAACTTCATG 1694

; RESULT 9
; US-09-887-586A-57
; Sequence 57, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 57
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1983)
; OTHER INFORMATION: (-)-4S-limonene synthase
; US-09-903-012-57

Query Match      83.4%; Score 112.6; DB 9; Length 2089;
Best Local Similarity 89.6%; Pred. No. 6.5e-29;
Matches 121; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      1  CTTGATTACGAGGTGATACGCGGTCTACAAGCGGACAGGGCTCGTGAGAGAAGCT 60
Db      1654  CTTGACTACGAGGTGACACGCGGTCTACAAGCGGATAGGGCCCGTGGAGAGAAGCT 1713

QY      61  TCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 120
Db      1714  TCAGCTATATCGTGTATATGAAGACCATCTCTGGATCAATAGAGGAAGATGCTCTCAAT 1773

QY      121  CATATCAACGCCCATG 135
Db      1774  CATATCAACGCCCATG 1788

; RESULT 10
; US-09-903-012-57
; Sequence 57, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 57
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1983)
; OTHER INFORMATION: (-)-4S-limonene synthase
; US-09-903-012-57

Query Match      83.4%; Score 112.6; DB 9; Length 2089;
Best Local Similarity 89.6%; Pred. No. 6.5e-29;
Matches 121; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      1  CTTGATTACGAGGTGATACGCGGTCTACAAGCGGACAGGGCTCGTGAGAGAAGCT 60
Db      1654  CTTGACTACGAGGTGACACGCGGTCTACAAGCGGATAGGGCCCGTGGAGAGAAGCT 1713

QY      61  TCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTCGAT 120
Db      1714  TCAGCTATATCGTGTATATGAAGACCATCTCTGGATCAATAGAGGAAGATGCTCTCAAT 1773

QY      121  CATATCAACGCCCATG 135
Db      1774  CATATCAACGCCCATG 1788

; RESULT 11
; US-09-900-797-57
; Sequence 57, Application US/09900797
```

```
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 57
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1983)
; OTHER INFORMATION: (-)-4S-limonene synthase
US-09-900-797-57

Query Match      83.4%; Score 112.6; DB 10; Length 2089;
Best Local Similarity 89.6%; Pred. No. 6.5e-29;
Matches 121; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 CTTGCTATACGAGGTGATACGGCGTCTACAAAGCGGACAGGGCTCGTGGAGAAAGCT 60
Db 1654 CTTGCTATACGAGGTGATACGGCGTCTACAAAGCGGACAGGGCTCGTGGAGAAAGCT 1713
QY 61 TCCTCTATATCATGTTATATGAAGACAAATCCTGGAGTATCAGAGGAGATGCTCTCGAT 120
Db 1714 TCAGCTATATCGTGTATATGAAGACCAATCCTGGATCAATAGAGGAAGATGCTCTCAAT 1773
QY 121 CATATCAACGCCATG 135
Db 1774 CATATCAACGCCATG 1788

RESULT 13
US-10-041-007-23
; Sequence 23, Application US/10041007
; Publication No. US20020164736A1
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seiichi P.T.
; APPLICANT: Schepmann, Hala G.
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase
; FILE REFERENCE: P02081US1
; CURRENT APPLICATION NUMBER: US/10/041,007
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/259,881
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 23
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Abies grandis
US-10-041-007-23

Query Match      83.4%; Score 112.6; DB 14; Length 2089;
Best Local Similarity 89.6%; Pred. No. 6.5e-29;
Matches 121; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 CTTGCTATACGAGGTGATACGGCGTCTACAAAGCGGACAGGGCTCGTGGAGAAAGCT 60
Db 1654 CTTGCTATACGAGGTGATACGGCGTCTACAAAGCGGACAGGGCTCGTGGAGAAAGCT 1713
QY 61 TCCTCTATATCATGTTATATGAAGACAAATCCTGGAGTATCAGAGGAGATGCTCTCGAT 120
Db 1714 TCAGCTATATCGTGTATATGAAGACCAATCCTGGATCAATAGAGGAAGATGCTCTCAAT 1773
QY 121 CATATCAACGCCATG 135
Db 1774 CATATCAACGCCATG 1788

RESULT 14
US-10-025-145A-5
; Sequence 5, Application US/10025145A
; Publication No. US20030175861A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Bohlmann, Joerg
; APPLICANT: Steele, Christopher L.
; APPLICANT: Phillips, Michael A.
; TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)
; FILE REFERENCE: WSUR118414
; CURRENT APPLICATION NUMBER: US/10/025,145A
; CURRENT FILING DATE: 2002-06-28
```

Fri Jul 9 08:49:07 2004

Search completed: July 8, 2004, 17:14:28  
Job time : 71.7053 secs

```

; PRIOR APPLICATION NUMBER: US 09/360,545
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: PCT/US98/14528
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: US 60/052,249
; PRIOR FILING DATE: 1997-07-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Abies Grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)..(1986)
; OTHER INFORMATION:
US-10-025-145A-5

Query Match      83.4%; Score 112.6; DB 15; Length 2089;
Best Local Similarity 89.6%; Pred. No. 6.5e-29;
Matches 121; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      1  CTTCCGATTACGAGGTGATACGGGTGCTACAGCGCGGACAGGGCTCGTGGAGAGAGCT 60
Db      1654 CTTCCGACTACGAGGTGACACGGCGTCTCTACAGGGCGGATAGGGCCGCTGGAGAGAGCT 1713

QY      61  TCCTCTATATCATGTTATATGAAGACAACTCTGGAGTATCAGAGGAGATGCTCTCGAT 120
Db      1714 TCAGCTATATCGTGTATATGAAGACCACTCTGGATCAATAGAGGAGAGATGCTCTCAAT 1773

QY      121 CATATCAAGCCCATG 135
Db      1774 CATATCAAGCCCATG 1788

RESULT 15
US-10-041-007-27
; Sequence 27, Application US/10041007
; Publication No. US20020164736A1
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seiichi P.T.
; APPLICANT: Schepmann, Hala G
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase
; FILE REFERENCE: P02081US1
; CURRENT APPLICATION NUMBER: US/10/041,007
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/259,881
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 2429
; TYPE: DNA
; ORGANISM: Abies grandis
US-10-041-007-27

Query Match      82.2%; Score 111; DB 14; Length 2429;
Best Local Similarity 88.9%; Pred. No. 2.5e-28;
Matches 120; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1  CTTCCGATTACGAGGTGATACGGGTGCTACAGCGCGGACAGGGCTCGTGGAGAGAGCT 60
Db      1616 CTTCCGGTACCTGTTGACACGGCTGCTACAGGGCGGATAGGGCTCGTGGAGAGAGCT 1675

QY      61  TCCTCTATATCATGTTATATGAAGACAACTCTGGAGTATCAGAGGAGATGCTCTCGAT 120
Db      1676 TCAGCTATATCGTGTATATGAAGACCACTCTGGATCAATAGAGGAGAGATGCTCTCAAT 1735

QY      121 CATATCAAGCCCATG 135
Db      1736 CATATCAAGCCCATG 1750

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2004, 11:16:49 ; Search time 413.736 Seconds  
(without alignments)  
9743.880 Million cell updates/sec

Title: US-10-025-145A-3\_COPY\_1560\_1694

Perfect score: 135  
Sequence: 1 cttgatgaagtgatgac.....tcgatcatatcaacgccatg 135

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estlin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_estc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_est3:\*

12: gb\_est4:\*

13: gb\_est5:\*

14: gb\_est6:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vit:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vri:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115.8	85.8	516	13	BQ698077
2	115.8	85.8	541	14	CF672979
3	115.8	85.8	637	14	CF476978
4	115.8	85.8	697	13	BQ196773

5	115.8	85.8	745	14	CF477103
6	112.6	83.4	517	13	BX677624
7	111.8	82.8	599	12	BG526917
8	104.6	77.5	591	14	CA305371
9	104.6	77.5	599	13	BX680641
10	96.6	71.6	542	10	BG039521
11	96.6	71.6	574	9	AW065088
12	96.6	71.6	574	14	CF474640
13	96.6	71.6	651	14	CF479884
14	96.6	71.6	696	14	CF401916
15	96.6	71.6	700	14	CF474786
16	96.6	71.6	733	14	CF666270
17	96.6	71.6	740	14	CF477562
18	88.6	65.6	597	14	CF666416
19	88.6	65.6	682	14	CF663768
20	74.2	55.0	669	14	CF479802
21	66.4	49.2	539	10	AW287756
22	61.4	45.5	534	14	CF389058
23	61.4	45.5	741	14	CF389118
24	59	43.7	488	14	CF472708
25	59	43.7	669	14	CF397861
26	59	43.7	682	14	CF397245
27	55.6	41.2	635	14	CF665839
28	51	37.8	400	14	CF672970
29	49.8	36.9	730	13	BX682869
30	49.6	36.7	506	12	BG317597
31	46.2	34.2	567	9	AU299970
32	40.6	30.1	515	14	CF663952
33	36.6	27.1	360	10	AW290355
34	36.6	27.1	360	14	CD026983
35	36.2	26.8	485	14	CF470359
36	35.8	26.5	990	29	CG014635
37	35.6	26.4	380	29	CG999639
38	34.8	25.8	414	29	CG757577
39	34.8	25.8	619	29	CC826228
40	34.8	25.8	633	29	CG848835
41	34.8	25.8	688	29	CC875261
42	34.8	25.8	745	29	CC740342
43	34.8	25.8	799	29	CG884155
44	34.6	25.6	690	12	BG591505
45	34.2	25.3	540	29	CG888779

## ALIGNMENTS

RESULT 1

BQ698077

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

BQ698077 516 bp mRNA linear EST 07-MAY-2003  
NXPV\_064\_C05\_F NXPV (Nsf Xylem Planings wood Vertical) Pinus taeda  
cDNA clone NXPV\_064\_C05\_5' similar to Arabidopsis thaliana sequence  
At1G61680 hypofunctional protein see  
<http://mips.gsf.de/proj/thal/db/index.html>, mRNA sequence.

BQ698077 GI:21823393

EST

Pinus taeda (loblolly pine)

Pinus taeda

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
1 (bases 1 to 516)

Sederoff, R.

Molecular Basis of Wood Formation in the Pine Megagenome

Unpublished (2000)

Contact: Sederoff, Ron

Forest Biotechnology

North Carolina State University

840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,

NC 27695, USA

Tel: 919 515 7800

Fax: 919 515 7801

Email: [ron.sederoff@ncsu.edu](mailto:ron.sederoff@ncsu.edu), [jerry.johnson@ncsu.edu](mailto:jerry.johnson@ncsu.edu)

Please see <http://web.ahc.umn.edu/biodata/nsfpine/> for further

information.

Seq primer: T3  
source

# FEATURES

Location/Qualifiers  
1..516  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="Coastal plain loblolly pine from North Carolina"  
/db\_xref="taxon:3152"  
/clone="NXPV\_064 C05"  
/tissue\_type="Xylem"  
/cell\_type="Planings (secondary)"  
/dev\_stage="Transitional"  
/lab\_host="X11-Blue"  
/clone\_lib="NXPV (Nsf Xylem Planings wood Vertical)"  
/note="Vector: Bluescript SK; Site\_1: Eco RI; Site\_2: XhoI; The library is from early (spring) secondary wood, taken from a ten year old tree in the transitional phase. The tree is a kind gift of the Westvaco Corporation. Secondary xylem was harvested from the tree by peeling back the bark and primary xylem and then removing the underlying tissue with a block plane. NOTE: The sequences contain a 'cDNA adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'AATTCGGACGAG'."

# ORIGIN

Query Match 85.8%; Score 115.8; DB 13; Length 516;  
Best Local Similarity 91.1%; Pred. No. 4.6e-27;  
Matches 123; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 CTTTCGATTACGAGGTGATACGGCGGTCTACAGCGGACAGAGGCTCGTGGAGGAAGCT 60  
|||||  
DB 209 CTTTCGATTACGAGGATACGGCTCTACAGCGGACAGAGGCTCGTGGAGGAAGCT 268  
|||||

QY 61 TCCTCTATATCATGTTATATGAAGACAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 120  
|||||  
DB 269 TCGTGTATATCTGTTATATGAAGACAATCTCGAGTATCAGAGGAAGATGCTCTCAAT 328  
|||||

QY 121 CATATCAAGCCCATG 135  
|||||  
DB 329 CATCTCAAGCCCATG 343  
|||||

RESULT 2  
CF672979  
LOCUS  
DEFINITION  
R1C11\_75\_F01\_b1\_A029 Root control Pinus taeda cDNA clone  
CF672979 541 bp mRNA linear EST 07-OCT-2003  
CF672979\_R1C11\_75\_F01\_A029 3', mRNA sequence.  
CF672979.1 GI:37570372  
EST.  
Pinus taeda (loblolly pine)  
Pinus taeda  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
1 (bases 1 to 541)  
Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Zimmermann, C. and  
Dean, J.F.D.  
An EST database from untreated loblolly pine (Pinus taeda) roots  
Unpublished (2003)  
Other ESTs: R1C11\_75\_F01.g1\_A029  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
RNA prepared and library constructed by W. Walter Lorenz (School of  
Forest Resources, University of Georgia); plant material prepared  
by Craig Zimmermann (School of Forest Resources, University of  
Georgia) using rooted cuttings provided by the Forest Biology  
Research Cooperative (FBRC) and the CCLONES project a the

University of Florida; sequencing done in the Laboratory for  
Genomics and Bioinformatics, University of Georgia. Sequence ends  
have been trimmed to exclude vector and regions below phred quality  
16. Three-prime sequences are presented as their reverse complement  
and have been trimmed to exclude polyA.  
Seq primer: M13-21 (TGTAACACGACGGCCAGT)  
POLYA=Yes.

# FEATURES

Location/Qualifiers  
1..541  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="3 CCLONES"  
/db\_xref="taxon:3152"  
/clone="R1C11\_75\_F01\_A029"  
/lab\_host="DH10B-Ti phage-resistant E. coli"  
/clone\_lib="Root control"  
/note="Organ: root; Vector: pSL1180; Site\_1: EcoRI;  
Site\_2: XhoI; The library was prepared from polyA+ RNA  
from the roots of 1-year-old loblolly pine (Pinus taeda)  
cuttings that were rooted and then planted in washed sand.  
Just before harvesting roots for RNA isolation, the rooted  
cuttings were maintained for 27 days (April 2003) under  
ambient conditions in a local greenhouse. They were kept  
on a weekly regimen of 0.5x nutrient-complete Hoagland's  
solution and supplemented with additional water sufficient  
to maintain a 15% soil moisture content. Double-stranded  
cDNA was cloned unidirectionally into pSL1180. Inserts can  
be excised with EcoRI (5' end) and XhoI (3' end)."

# ORIGIN

Query Match 85.8%; Score 115.8; DB 14; Length 541;  
Best Local Similarity 91.1%; Pred. No. 4.7e-27;  
Matches 123; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 CTTTCGATTACGAGGTGATACGGCGGTCTACAGCGGACAGAGGCTCGTGGAGGAAGCT 60  
|||||  
DB 66 CTTTCGATTACGAGGATACGGCTCTACAGCGGACAGAGGCTCGTGGAGGAAGCT 125  
|||||

QY 61 TCCTCTATATCATGTTATATGAAGACAATCTCGAGTATCAGAGGAAGATGCTCTCGAT 120  
|||||  
DB 126 TCGTGTATATCTGTTATATGAAGACAATCTCGAGTATCAGAGGAAGATGCTCTCAAT 185  
|||||

QY 121 CATATCAAGCCCATG 135  
|||||  
DB 186 CATCTCAAGCCCATG 200  
|||||

# RESULT 3

CF476978  
LOCUS  
DEFINITION  
R1C11\_75\_F01\_b1\_A022 Well-watered loblolly pine roots WW3  
CF476978 637 bp mRNA linear EST 08-SEP-2003  
CF476978\_R1C11\_75\_F01\_A022 3', mRNA sequence.  
CF476978.1 GI:34505847  
EST.  
Pinus taeda (loblolly pine)  
Pinus taeda  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
1 (bases 1 to 637)  
Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J.,  
Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and  
Neale, D.  
An EST database from well-watered loblolly pine (Pinus taeda) roots  
Unpublished (2003)  
Other ESTs: R1C11\_75\_F01.g1\_A022  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.

Seq primer: M13-21 (TGTAAACGACGGCCAGT)  
POLYA=Yes.

## FEATURES

source

Location/Qualifiers

1..637  
/organism="Pinus taeda"  
/mol\_type="rRNA"  
/strain="CCONES"  
/db\_xref="taxon:3352"  
/clone="RTW3\_5\_A06\_A022"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="Well-watered loblolly pine roots W3"  
/note="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

## ORIGIN

Query Match 85.8%; Score 115.8; DB 14; Length 637;  
Best Local Similarity 91.1%; Pred. No. 5.1e-27;  
Matches 123; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 CTTGATTACGAGGTGATACCGGTGCTACAGGCGGACAGGGCTCGTGGAGAGAAAGCT 60  
|||||  
Db 225 CTTGATTACGAGGAGATACCGGTGCTACCGAGCGGACAGGGCCCGTGGAGAGAAAGCT 284  
|||||

Qy 61 TCCTCTATATCATGTTATATGAAGACAAATCTGGAGTATCAGAGGAAGATGCTCTCGAT 120  
|||||  
Db 285 TCGGTATATCTTGTATATGAAGACAAATCTGGAGCAACGAGAGAAAGTCTCTCAAT 344  
|||||

Qy 121 CATATCAACGCCATG 135  
Db 345 CATATCAACGCCATG 359  
|||||

## RESULT 4

BQ196773

LOCUS

DEFINITION

BQ196773 697 bp mRNA linear EST 07-MAY-2003  
NXIV105\_B02\_F NXIV (Nsf Xylem Late wood Vertical) Pinus taeda cDNA  
clone NXIV105\_B02\_5' similar to Arabidopsis thaliana sequence  
At4g16730 linonene cyclase like protein see  
http://mips.gsf.de/proj/thal/db/index.html, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

COMMENT

BQ196773.1 GI:20379276

EST.

Pinus taeda (loblolly pine)

Pinus taeda

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

1 (bases 1 to 697)

Sederoff, R.

Molecular Basis of Wood Formation in the Pine Megagenome

Unpublished (2000)

Contact: Sederoff, Ron

Forest Biotechnology

North Carolina State University

840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,

NC 27695, USA

Tel: 919 515 7800

Fax: 919 515 7801

Email: ron.sederoff@ncsu.edu, jerri.johnson@ncsu.edu

Please see http://web.ahc.umn.edu/biodata/nsfpine/ for further

information.

Seq primer: T3.

## FEATURES

source

Location/Qualifiers

1..697  
/organism="Pinus taeda"  
/mol\_type="rRNA"  
/strain="Coastal plain loblolly pine from North Carolina"  
/db\_xref="taxon:3352"  
/clone="NXIV105\_B02"  
/tissue\_type="primary xylem"  
/dev\_stage="late wood"  
/lab\_host="XLI-Blue"  
/clone\_lib="NXIV (Nsf Xylem Late wood Vertical)"  
/note="Vector: pTriplex; Site 1: EcoRI; The library is from late (summer-August) wood, taken from below the crown of a 20 year old tree. The harvested xylem tissue was on the cusp between transitional and mature wood. NOTE: The sequences contain a 'cDNA adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'AATCGCCATTATGCCC'."

## ORIGIN

Query Match 85.8%; Score 115.8; DB 13; Length 697;  
Best Local Similarity 91.1%; Pred. No. 5.3e-27;  
Matches 123; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 CTTGATTACGAGGTGATACCGGTGCTACAGGCGGACAGGGCTCGTGGAGAGAAAGCT 60  
|||||  
Db 332 CTTGATTACGAGGAGATACCGGTGCTACCGAGCGGACAGGGCCCGTGGAGAGAAAGCT 391  
|||||

Qy 61 TCCTCTATATCATGTTATATGAAGACAAATCTGGAGTATCAGAGGAAGATGCTCTCGAT 120  
|||||  
Db 392 TCGGTATATCTTGTATATGAAGACAAATCTGGAGCAACAGAGAGAGATGCTCTCAAT 451  
|||||

Qy 121 CATATCAACGCCATG 135  
Db 452 CATCTCAACGCCATG 466  
|||||

## RESULT 5

CF477103

LOCUS

DEFINITION

CF477103 745 bp mRNA linear EST 08-SEP-2003  
RTW3\_5\_A06\_g1\_A022 Well-watered loblolly pine roots WW3 Pinus  
taeda cDNA clone RTW3\_5\_A06\_A022 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CF477103.1 GI:34505972

EST.

Pinus taeda (loblolly pine)

Pinus taeda

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

1 (bases 1 to 745)

Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and Neale, D.

An EST database from well-watered loblolly pine (Pinus taeda) roots

Unpublished (2003)

Other ESTs: RTW3\_5\_A06.bl\_A022

Contact: Cordonnier-Pratt, M

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz, School of

Forestry, University of Georgia; plant material prepared at the

University of Florida; sequencing done in the Laboratory for

Genomics and Bioinformatics, University of Georgia. Sequence ends

have been trimmed to exclude vector and regions below Phred quality

16. Three-prime sequences are presented as their reverse complement

and have been trimmed to exclude polyA.

Seq primer: JENREV (CAGGAACAGCTATGACC).

Location/Qualifiers

1..745

```

/organism="Pinus taeda"
/mol_type="mRNA"
/strain="CCLONES"
/db_xref="taxon:3352"
/clone="RTW3_5_A06_A022"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="Well-watered loblolly pine roots W3"
/note="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."
```

## ORIGIN

```

Query Match      85.8%; Score 115.8; DB 14; Length 745;
Best Local Similarity 91.1%; Pred. No. 5.5e-27;
Matches 123; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 CTTTCGATTACGAGGTGATACGGCGTCTACAGCGGACAGGGCTCGTGAGAGAAGCT 60
D6 461 CTTTCGATTACGAGGATACGCGCTGCTACAGCGGACAGGGCGCCGCTGGAGAAGAGCT 520
QY 61 TCCTCTATATCATCTGTTATATGAAGACAATCCTCGAGTATCAGAGGAAGATGCTCTCGAT 120
D6 521 TCCTGTATATCTGTTATATGAAGACAATCCTGGAGCAACGGAGAGATGCTCTCAAT 580
QY 121 CATATCAACGCCATG 135
D6 581 CATATCAACGCCATG 595
```

```

RESULT 6
BX677624
LOCUS BX677624 RN Pinus pinaster cDNA clone RN42B08, mRNA sequence.
DEFINITION BX677624
ACCESSION BX677624
VERSION BX677624.1 GI:38011576
KEYWORDS EST.
SOURCE Pinus pinaster
Pinus pinaster
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
Frigerio, J. and Plomion, C.
1 (bases 1 to 517)
Identification of water-deficit responsive genes in Maritime pine
(Pinus pinaster Ait.) using an EST approach
Unpublished (2002)
Contact: Frigerio JM
Genetique et Amelioration 69
INRA
route d'Arcachon 33612 Cestas CEDEX France
Email: Frigerio@pierrot.inra.fr
Email: Frigerio@pierrot.inra.fr
Seq primer: T3.
```

## FEATURES

```

source
1..517
Location/Qualifiers
/organism="Pinus pinaster"
/mol_type="mRNA"
/db_xref="taxon:71647"
/clone="RN42B08"
/tissue_type="root"
/dev_stage="6 weeks old seedling"
/lab_host="SOLR"
/clone_lib="RN"
```

```

/note="vector: Uni-ZAP XR; ecotype: Landes; The library was made from the roots of 6 weeks old seedlings grown in hydroponic conditions. A mixture of genotypes were used. Oligo-dT primed cDNA was directionally cloned into the EcoRI-XhoI lambda-ZAP vector arms and mass-excised to form a pBluescript phagemid"
```

## ORIGIN

```

Query Match      83.4%; Score 112.6; DB 13; Length 517;
Best Local Similarity 89.6%; Pred. No. 5.2e-26;
Matches 121; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 CTTTCGATTACGAGGTGATACGGCGTCTACAGCGGACAGGGCTCGTGAGAGAAGCT 60
D6 345 CTTTCGATTACGAGGCGACACTGCTCTACAGCGGACAGGGCGCCGCTGGAGAAGAGCT 404
QY 61 TCCTCTATATCATCTGTTATATGAAGACAATCCTCGAGTATCAGAGGAAGATGCTCTCGAT 120
D6 405 TCGGGTATATCTGTTATATGAAGACAATCCTCGAACAACACAGAGGAAGATGCTCTCAAT 464
QY 121 CATATCAACGCCATG 135
D6 465 CATATCAACGCCATG 479
```

## RESULT 7

```

BG526917
LOCUS BG526917
DEFINITION BG526917 599 bp mRNA linear EST 07-MAY-2003
NXPV_057_D04_F_NXPV (Nsf Xylem Planings wood Vertical) Pinus taeda
cDNA clone NXPV_057_D04_5' similar to Arabidopsis thaliana sequence
At4g16730 limonene cyclase like protein see
http://mips.gsf.de/proj/thal/db/index.html, mRNA sequence.
BG526917
ACCESSION BG526917.1 GI:13536796
VERSION
KEYWORDS EST.
SOURCE Pinus taeda (loblolly pine)
Pinus taeda
```

## ORGANISM

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
Sederoff, R.
```

## REFERENCE

```

1 (bases 1 to 599)
Molecular Basis of Wood Formation in the Pine Megagenome
Unpublished (2000)
Contact: Sederoff, Ron
Forest Biotechnology
North Carolina State University
840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,
NC 27695, USA
Tel: 919 515 7800
Fax: 919 515 7801
Email: ron.sederoff@ncsu.edu, jerri.johnson@ncsu.edu
Please see http://web.ahc.umn.edu/biodata/nsfpine/ for further
information.
```

Seq primer: T3.

Location/Qualifiers

## FEATURES

## source

```

1..599
/organism="Pinus taeda"
/mol_type="mRNA"
/strain="Coastal plain loblolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="NXPV_057_D04"
/tissue_type="Xylem"
/cell_type="Planings (secondary)"
/dev_stage="Transitional"
/lab_host="XLI-Blue"
/clone_lib="NXPV (Nsf Xylem Planings wood Vertical)"
/note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI; The library is from early (spring) secondary wood, taken from a ten year old tree in the transitional phase. The tree is a kind gift of the Westvaco Corporation. Secondary xylem was harvested from the tree by peeling back the bark and primary xylem and then removing the underlying tissue with a block plane. NOTE: The sequences contain a 'cDNA adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'ATTCCGACGAG'."
```

## ORIGIN

```

Query Match      82.8%; Score 111.8; DB 12; Length 599;
Best Local Similarity 86.1%; Pred. No. 1e-25;
```

Matches 119; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 CTTTCATTACGAGGTGATACGGGTGCTACAGGCGGACAGGCGTCTGGAGAGAAGCT 60  
 |||||  
 Db 345 CTTTCATTACGAGGATACGGGTGCTACCCNNGCGGACAGGCGCTGGAGAGAAGCT 404  
 |||||

QY 61 TCCTCTATATCATGTTATATGAAGACATCTCTGGGTATATCAGAGGAAGATCTCTCGAT 120  
 |||||  
 Db 405 TCGTGTTATCTGTTATATGAAGACATCTCTNNAACACAGAGGAAGATCTCTCAAT 464  
 |||||

QY 121 CATATCAAGCCCATG 135  
 |||||  
 Db 465 CATCTCAAGCCCATG 479  
 |||||

RESULT 8  
 CA305371  
 LOCUS  
 DEFINITION  
 hasp004xk08 Heterobasidion annosum - Scots pine infection stage  
 subtraction cDNA library (hasp) Pinus sylvestris/Heterobasidion  
 annosum cDNA clone hasp004xk08, mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Pinus sylvestris/Heterobasidion annosum  
 Pinus sylvestris/Heterobasidion annosum  
 Eukaryota; mixed EST libraries.

REFERENCE  
 1 (bases 1 to 591)  
 Asiegbu, F.O., Nahalkova, J., Choi, W., Stenlid, J., and Dean, R.A.  
 Expressed sequence tags of randomly selected cDNA clones from the  
 interaction of the root rot fungus (Heterobasidion annosum) with  
 seedling roots of Scots pine (Pinus sylvestris)  
 Unpublished (2001)

JOURNAL  
 COMMENT  
 Contact: Fred O. Asiegbu  
 Dept. of Forest Mycology & Pathology  
 Swedish University of Agriculture, Box 7026, S-750 07, Uppsala,  
 Sweden  
 Tel: +46 18 67 15 98  
 Fax: +46 18 30 92 45  
 Email: Fred.Asiegbu@mykopat.slu.se  
 Seq primer: T7 primer.

FEATURES  
 source  
 1..591  
 Location/Qualifiers  
 /organism="Pinus sylvestris/Heterobasidion annosum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:169015"  
 /clone="hasp004xk08"  
 /dev\_stage="Seedling roots of scots pine were infected for  
 6 days with H. annosum"  
 /clone\_lib="Heterobasidion annosum - Scots pine infection  
 stage subtraction cDNA library (hasp)"  
 /note="vector: pT-Adv; Site\_1: EcoRI; The subtractive  
 hybridization cDNA library was constructed from scots  
 pine roots infected for 6-days with mycelia of  
 Heterobasidion annosum (FP5)."

ORIGIN  
 Query Match 77.5%; Score 104.6; DB 14; Length 591;  
 Best Local Similarity 85.9%; Pred. No. 2.3e-23;  
 Matches 116; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 CTTTCATTACGAGGTGATACGGGTGCTACAGGCGGACAGGCGTCTGGAGAGAAGCT 60  
 |||||  
 Db 163 CTTTCATTACGAGGTGATACGGGTGTTACAGGCGGACAGTGTCTGGAGAGAAGCT 222  
 |||||

QY 61 TCCTCTATATCATGTTATATGAAGACATCTCTGGGTATATCAGAGGAAGATCTCTCGAT 120  
 |||||  
 Db 223 TCGTGTTATCTGTTATATGAAGACATCTCTNNAACACAGAGGAAGATCTCTCAAT 282  
 |||||

QY 121 CATATCAAGCCCATG 135  
 |||||  
 Db 283 CATATCAAGCCCATG 297  
 |||||

RESULT 9  
 BX680641  
 LOCUS  
 DEFINITION  
 RS Pinus pinaster cDNA clone RS46D06, mRNA sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Pinus pinaster  
 Pinus pinaster  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

REFERENCE  
 1 (bases 1 to 599)  
 Frigerio, J. and Plomion, C.  
 Identification of water-deficit responsive genes in Maritime pine  
 (Pinus pinaster Ait.) using an EST approach  
 Unpublished (2002)  
 Contact: Frigerio JM  
 Genetique et Amelioration 69  
 INRA  
 route d'Arcachon 33612 Cestas CEDEX France  
 Email: Frigerio@pierrot.inra.fr  
 Email: Frigerio@pierrot.inra.fr  
 Seq primer: T3.

FEATURES  
 source  
 1..599  
 Location/Qualifiers  
 /organism="Pinus pinaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:71647"  
 /clone="RS46D06"  
 /tissue\_type="root"  
 /dev\_stage="6 weeks old seedling"  
 /lab\_host="SOLR"  
 /clone\_lib="RS"

/note="Vector: Uni-ZAP XR; ecotype: Landes; The library  
 was made from the roots of 6 weeks old seedlings grown in  
 hydroponic conditions. A three weeks drought stress  
 treatment was applied by lowering the osmotic potential of  
 the nutrient solution to -0.45 MPa using 170 g/l of  
 polyethylene glycol as an osmoticum. A mixture of  
 genotypes were used. Oligo-dT primed cDNA was  
 directionally cloned into the EcoRI-XhoI lambda-ZAP vector  
 arms and mass-excised to form a pBluescript phagemid"

ORIGIN  
 Query Match 77.5%; Score 104.6; DB 13; Length 599;  
 Best Local Similarity 85.9%; Pred. No. 2.3e-23;  
 Matches 116; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 CTTTCATTACGAGGTGATACGGGTGCTACAGGCGGACAGGCGTCTGGAGAGAAGCT 60  
 |||||  
 Db 332 CTTTCATTACGAGGTGATACGGGTGTTATCAGGAGGACAGGCGTCTGGAGAGAAGCT 391  
 |||||

QY 61 TCCTCTATATCATGTTATATGAAGACATCTCTGGGTATATCAGAGGAAGATCTCTCGAT 120  
 |||||  
 Db 392 TCGTGTTATCTGTTATATGAAGACATCTCTNNAACACAGAGGAAGATCTCTTAAT 451  
 |||||

QY 121 CATATCAAGCCCATG 135  
 |||||  
 Db 452 CATCTCAATGTCATG 466  
 |||||

RESULT 10  
 BG039521  
 LOCUS  
 DEFINITION  
 NXSI\_099\_H06\_F NXSI (Nsf Xylem Side wood Includ) Pinus taeda cDNA  
 clone NXSI\_099\_H06\_5, similar to Arabidopsis thaliana sequence  
 At1g66020 unknown protein see  
 http://mips.gsf.de/proj/thal/db/index.html, mRNA sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 BG039521  
 EST.  
 Pinus taeda (loblolly pine)

ORGANISM Pinus taeda  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
REFERENCE 1 (bases 1 to 542)  
AUTHORS Sederoff, R.  
TITLE Molecular Basis of Wood Formation in the Pine Megagenome  
JOURNAL Unpublished (2000)  
COMMENT Contact: Sederoff, Ron  
Forest Biotechnology  
North Carolina State University  
840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh, NC 27695, USA  
Tel: 919 515 7800  
Fax: 919 515 7801  
Email: ron.sederoff@ncsu.edu, jerri.johnson@ncsu.edu  
Please see <http://web.ahc.umn.edu/biodata/nsfpine/> for further information.  
Seq primer: T3.

FEATURES  
source  
1..542  
Location/Qualifiers  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="Coastal plain loblolly pine from North Carolina"  
/db\_xref="taxon:3352"  
/clone="NXSI\_099 H06"  
/tissue\_type="Xylem"  
/cell\_type="Side"  
/dev\_stage="Juvenile"  
/lab\_host="Xl1-Blue"  
/clone\_lib="NXSI (Nsf Xylem Side wood Inclined)"  
/note="Vector: Bluescript SK; Site 1: Eco RI; Site 2: XhoI; The library is from early (spring) wood, taken from three six-year old trees (three different genotypes), in the juvenile phase. These trees were induced to form side wood by bending to a 45 degree angle and tying them to the ground. Differentiating xylem was harvested from the sides of the inclined stems, and a mixture of all three genotypes was used for the library. oligo-dT primed cDNA was directionally cloned into the EcoRI-XhoI Bluescript SK vector arms. NOTE: The sequences contain a 'cDNA adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'AATTCGGCAGAG'."

ORIGIN  
Query Match 71.6%; Score 96.6; DB 10; Length 542;  
Best Local Similarity 82.2%; Pred. No. 9.5e-21;  
Matches 111; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
QY 1 CTTTCGATTACGAGTGATACGGGTGCTACAAGCGGACAGGGCTCGTGAGAGAAGCT 60  
Db 90 CTTTCGACTGAAGGTGACACTCAATGCTACAAGGCTGACAGGGCGGCTGGAGAAGAAGCT 149  
QY 61 TCCTCTATATCATGTTATATGAAGACAATCCTCGAGTATCAGAGGAGATGCTCTCGAT 120  
Db 150 TCGCGCGTATCGTGTATATGAAGACCACTCCTGGATATACAGAGGAAGATGCTGTCAAT 209  
QY 121 CATATCAACGCCATG 135  
Db 210 CAACTCAATGCTATG 224

RESULT 11  
LOCUS AW065088 574 bp mRNA linear EST 12-OCT-1999  
DEFINITION ST39E04 Pine Triplex shoot tip library Pinus taeda cDNA clone ST39E04, mRNA sequence.  
ACCESSION AW065088  
VERSION AW065088.1 GI:6020160  
KEYWORDS EST.  
SOURCE Pinus taeda (loblolly pine)  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

REFERENCE  
AUTHORS Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.  
TITLE The Pine Gene Discovery Project  
JOURNAL Unpublished (1999)  
COMMENT Contact: Ross Whetten

Forest Biotechnology Group  
North Carolina State University  
Dept. of Forestry, NC State University, 6113 Jordan Hall,  
Raleigh, NC, 27695-8008  
Tel: 919-515-7800  
Fax: 919-515-7801  
Email: rosswhetten@unity.ncsu.edu  
Seq primer: 5' lambda Triplex2 Sequencing Primer.

FEATURES  
source

1..574  
Location/Qualifiers  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/db\_xref="taxon:3352"  
/clone="ST39E04"  
/lab\_host="E. coli BM25.8"  
/note="Organ: shoot tips; Vector: Lambda Triplex; Site 1: SfiI (A); Site 2: SfiI (B); Shoot tips (approx. 2 cm from apex) were collected during the spring, frozen and used for mRNA isolation. The SMART-PCR method (Clontech) was used to prepare a library from 1 ug total RNA, using the Lambda Triplex vector. Plasmid subclones in pTriplex were recovered by cre-lox excision in E. coli strain BM25.8 and sequenced from the 5' end."

ORIGIN

Query Match 71.6%; Score 96.6; DB 9; Length 574;  
Best Local Similarity 82.2%; Pred. No. 9.7e-21;  
Matches 111; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
QY 1 CTTTCGATTACGAGTGATACGGGTGCTACAAGCGGACAGGGCTCGTGAGAGAAGCT 60  
Db 23 CTTTCGACTGAAGGTGACACTCAATGCTACAAGGCTGACAGGGCGGCTGGAGAAGAAGCT 82  
QY 61 TCCTCTATATCATGTTATATGAAGACAATCCTCGAGTATCAGAGGAAGATGCTCTCGAT 120  
Db 83 TCGCGCGTATCGTGTATATGAAGACCACTCCTGGATATACAGAGGAAGATGCTGTCAAT 142  
QY 121 CATATCAACGCCATG 135  
Db 143 CAACTCAATGCTATG 157

RESULT 12  
LOCUS CF474640 574 bp mRNA linear EST 05-SEP-2003  
DEFINITION RTW2\_7 B11\_b1 A021 Well-watered loblolly pine roots WW2 Pinus taeda cDNA clone RTW2\_7 B11\_A021 3', mRNA sequence.

ACCESSION CF474640  
VERSION CF474640.1 GI:34492012  
KEYWORDS EST.  
SOURCE Pinus taeda (loblolly pine)  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

REFERENCE 1 (bases 1 to 574)  
AUTHORS Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and Neale, D.  
TITLE An EST database from well-watered loblolly pine (Pinus taeda) roots Unpublished (2003)  
JOURNAL Other\_ESTs: RTW2\_7 B11\_g1\_A021  
COMMENT Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210

Email: mmpratt@uga.edu  
 RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
 Seq primer: M13-21 (TGTAACACGACGCCAGT)  
 POLYA=Yes.

#### FEATURES

Location/Qualifiers  
 1..574  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="CCIONES"  
 /db\_xref="taxon:3352"  
 /clone="RTW3\_12\_D02\_A021"  
 /lab\_host="DH10B-T1 phage-resistant E. coli"  
 /clone\_lib="Well-watered loblolly pine roots W3"  
 /note="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

#### ORIGIN

Query Match 71.6%; Score 96.6; DB 14; Length 574;  
 Best Local Similarity 82.2%; Pred. No. 9, 7e-21;  
 Matches 111; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1 CTTTCATACAGGTGATACCGGTGTCTACAAAGCGGACAGGGCTCGTGGAGAGAGCT 60  
 Db 50 CTTTCGACTGAAAGGTGACACTCAATGCTTACAAAGCTGACAGGGCGCGTGGAGAGAGCT 109  
 Qy 61 TCCTCTATATCATCTTATATGAAGACCACTCTCGAGTATCAGAGGAGATGCTCTCGAT 120  
 Db 110 TCGCCCGTATCGTGTATATGAAGACCACTCTCGAATACAGAGGAGATGCTGTCAAT 169  
 Qy 121 CATATCAACGCCATG 135  
 Db 170 CAAGTCAATGCTATG 184

RESULT 13  
 CF479884/c  
 LOCUS  
 DEFINITION  
 CF479884 651 bp mRNA linear EST 08-SEP-2003  
 taeda cDNA clone RTW3\_12\_D02\_A022 5', mRNA sequence.  
 CF479884  
 EST.  
 SOURCE  
 ORGANISM  
 Pinus taeda (loblolly pine)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
 1 (Bases 1 to 651)  
 Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J., Cannon, R., Owen, A. and Neale, D.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 EST database from well-watered loblolly pine (Pinus taeda) roots  
 Unpublished (2003)  
 Other ESTs: RTW3\_12\_D02.b1\_A022  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the

University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
 Seq primer: JENREV (CAGGAACAGCTATGACC).  
 Location/Qualifiers  
 1..651  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="CCIONES"  
 /db\_xref="taxon:3352"  
 /clone="RTW3\_12\_D02\_A022"  
 /lab\_host="DH10B-T1 phage-resistant E. coli"  
 /clone\_lib="Well-watered loblolly pine roots W3"  
 /note="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

#### FEATURES

Location/Qualifiers  
 1..651  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="CCIONES"  
 /db\_xref="taxon:3352"  
 /clone="RTW3\_12\_D02\_A022"  
 /lab\_host="DH10B-T1 phage-resistant E. coli"  
 /clone\_lib="Well-watered loblolly pine roots W3"  
 /note="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

#### ORIGIN

Query Match 71.6%; Score 96.6; DB 14; Length 651;  
 Best Local Similarity 82.2%; Pred. No. 1e-20;  
 Matches 111; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1 CTTTCATACAGGTGATACCGGTGTCTACAAAGCGGACAGGGCTCGTGGAGAGAGCT 60  
 Db 461 CTTTCGACTGAAAGGTGACACTCAATGCTTACAAAGCTGACAGGGCGCGTGGAGAGAGCT 402  
 Qy 61 TCCTCTATATCATCTTATATGAAGACCACTCTCGAGTATCAGAGGAGATGCTCTCGAT 120  
 Db 401 TCGCCCGTATCGTGTATATGAAGACCACTCTCGAATACAGAGGAGATGCTGTCAAT 342  
 Qy 121 CATATCAACGCCATG 135  
 Db 341 CAAGTCAATGCTATG 327

RESULT 14  
 CF401916  
 LOCUS  
 DEFINITION  
 CF401916 696 bp mRNA linear EST 29-AUG-2003  
 taeda cDNA clone RTW1\_15\_B05\_A015 5', mRNA sequence.  
 CF401916  
 EST.  
 SOURCE  
 ORGANISM  
 Pinus taeda (loblolly pine)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
 1 (Bases 1 to 696)  
 Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and Neale, D.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 An EST database from well-watered loblolly pine (Pinus taeda) roots  
 Unpublished (2003)  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2004, 09:55:09 ; Search time 8428.74 Seconds  
(without alignments)  
10351.438 Million cell updates/sec

Title: US-10-025-145A-64  
Perfect score: 2013  
Sequence: 1 ttttgagtgctcttctatc.....aaaaaaaaaaaaaaaaaaaaa 2013

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vit.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2013	100.0	2013	6	AR222136	AR222136 Sequence
2	2013	100.0	2013	8	AGU87910	U87910 Abies grand
3	1306.8	64.9	2018	6	BD227661	BD227661 Syntheses
4	1306.8	64.9	2018	6	AR222097	AR222097 Sequence
5	1306.8	64.9	2018	6	AR240696	AR240696 Sequence
6	1306.8	64.9	2018	6	AR266971	AR266971 Sequence
7	1306.8	64.9	2018	6	AR316320	AR316320 Sequence
8	1306.8	64.9	2018	6	AR338463	AR338463 Sequence
9	1306.8	64.9	2018	6	AR429869	AR429869 Sequence
10	1306.8	64.9	2018	8	AGU87909	U87909 Abies grand
11	1273.6	63.3	1958	8	AY237645	AY237645 Picea sit
12	1178.4	58.5	2082	8	AF543527	AF543527 Pinus tae
13	1155.6	57.4	2162	8	AF543529	AF543529 Pinus tae
14	1103.2	54.8	2186	8	AF461460	AF461460 Picea abi
15	1101.4	54.7	2198	8	AF369918	AF369918 Picea abi
16	1095.4	54.4	1893	8	AF139206	AF139206 Abies gra
17	1092.4	54.3	1890	6	AR222146	AR222146 Sequence
18	1085.8	53.9	2150	8	AF369919	AF369919 Picea abi
19	1072.4	53.3	2186	6	AR222137	AR222137 Sequence
20	1072.4	53.3	2186	8	AF139205	AF139205 Abies gra
21	1071.6	53.2	2196	6	BD227666	BD227666 Syntheses
22	1071.6	53.2	2196	6	AR222096	AR222096 Sequence
23	1071.6	53.2	2196	6	AR240701	AR240701 Sequence
24	1071.6	53.2	2196	6	AR266976	AR266976 Sequence
25	1071.6	53.2	2196	6	AR316325	AR316325 Sequence
26	1071.6	53.2	2196	6	AR338468	AR338468 Sequence
27	1071.6	53.2	2196	6	AR429874	AR429874 Sequence
28	1071.6	53.2	2196	8	AGU87908	U87908 Abies grand
29	1071.6	53.2	2205	6	AR222116	AR222116 Sequence
30	1040.2	51.7	2100	8	AF543530	AF543530 Pinus tae
31	971.4	48.3	1960	8	AF543531	AF543531 Pinus tae
32	929.4	46.2	2429	6	AR222138	AR222138 Sequence
33	929.4	46.2	2429	8	AF139207	AF139207 Abies gra
34	925.4	46.0	2089	6	BD227677	BD227677 Syntheses
35	925.4	46.0	2089	6	AR222098	AR222098 Sequence
36	925.4	46.0	2089	6	AR240712	AR240712 Sequence
37	925.4	46.0	2089	6	AR266987	AR266987 Sequence
38	925.4	46.0	2089	6	AR316336	AR316336 Sequence
39	925.4	46.0	2089	6	AR338479	AR338479 Sequence
40	925.4	46.0	2089	6	AR429885	AR429885 Sequence
41	925.4	46.0	2089	8	AF006193	AF006193 Abies gra
42	889.4	44.2	2084	8	AF543528	AF543528 Pinus tae
43	807.4	40.1	1513	6	BD272958	BD272958 Materials
44	805.8	40.0	1634	6	BD273051	BD273051 Materials
45	623.6	31.0	1173	6	BD273050	BD273050 Materials

ALIGNMENTS

RESULT 1	AR222136	Sequence 64	2013 bp	DNA	linear	PAT 26-SEP-2002
LOCUS	AR222136	Sequence 64	from patent US 6429014.			
DEFINITION	AR222136	Sequence 64	from patent US 6429014.			
ACCESSION	AR222136	Sequence 64	from patent US 6429014.			
VERSION	AR222136.1	GI:23329510				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 2013)					
AUTHORS	Steele,C.L., Bohlmann,J. and Croteau,R.B.					
TITLE	Monoterpene synthases from grand fir (Abies grandis)					
JOURNAL	Patent: US 6429014-A 64 06-AUG-2002;					
FEATURES	Location/Qualifiers					

[illegible]

RESULT 2	AGU87910	2013 bp	mRNA	linear	PLN 10-AUG-2001
LOCUS	Abies grandis (-)-camphene synthase (AG6.5)	mRNA	complete cds.		
DEFINITION	U87910				
ACCESSION	U87910.1	GI:2411484			
VERSION					
KEYWORDS					
SOURCE	Abies grandis				
ORGANISM	Abies grandis				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.					
1 (bases 1 to 2013)					
Bohlmann, J., Steele, C.L. and Croteau, R.					
Monoterpene synthases from grand fir ( <i>Abies grandis</i> ). cDNA isolation, characterization, and functional expression of myrcene synthase, (-)-(4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase					
J. Biol. Chem. 272 (35), 21784-21792 (1997)					
97413772					
9268308					
2 (bases 1 to 2013)					
Bohlmann, J., Phillips, M., Ramachandiran, V., Katoh, S. and Croteau, R.					
cDNA cloning, characterization, and functional expression of four new monoterpene synthase members of the Tpsd gene family from grand fir ( <i>Abies grandis</i> )					
Arch. Biochem. Biophys. 368 (2), 232-243 (1999)					
99373092					
10441373					
3 (bases 1 to 2013)					
Bohlmann, J., Steele, C.L. and Croteau, R.					
Direct Submission					
Submitted (31-JAN-1997)					
Institute of Biological Chemistry,					
Washington State University, Clark Hall, Pullman, WA 99164-6340,					
USA					
FEATURES	Location/Qualifiers				
1..2013	/organism="Abies grandis"				
/mol_type="mRNA"	/db_xref="taxon:46611"				
1..2013	/gene="AG6.5"				
36..1892	/note="terpene synthase; monoterpene synthase; terpene cyclase"				
/codon_start=1	/product="(-)-camphene synthase"				
/protein_id="AAB70707.1"	/db_xref="GI:2411485"				
/translation="MALLSITPLVSRCLSSSHEIKALRPIPLIGICRPGKSVASHI NMCLTSVASTDSVORVGNVHNSLWDDFIQSLISTPYGADPYRERADRLIGVKDIM FNFKLEGGNDLLQRLLLVDVBERLGDIDRHKKEIKTALDYVNSYNEKGIQGGRES VVTDLNSALGLRLTLRLHGYTVSSDLNVFKNGQFSSTANIQIEGIEIRGLVNLFRSLVAFPGKEWDEAETFTSKYLREALOKIPASSILSLEIRDLVYGHWTNLPRLERN YMDVFGQHTKNKNAAEKLLLELAENLFIHSLQREKLKHSRWKDSGSPMTFCRRH HVRYVALASCIAPFQPSGSGFLGTMKSHLITVLDDMYDVPGTVELELFTATIKRWD PSAMECLPEYMKGYMVHYHTVNMARVAEKQGRDITINVARQAEAFDSYMOEAKM IATGYLTFEYLENGKSSAHRPCALQPILTLDIPFPDHLIKEDVFPKINDLICII LRUGDTRCYKADRGAEESAISCYMKQNPGLTEEDALNHNFMIRDAIRELNWELL KPDSNVPIITSKKHAFDISRVVHHGYRDRDGYSPANVETKSLVMTVEIPVPL"					
Query Match	100.0%;	Score 2013;	DB 8;	Length 2013;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 2013;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	TTTTGAGCTGCTTCTTATCTGATAGCAAGCTGAAATGGCTCTCTTCTTCTATTACTCGCG	60		
Db	1	TTTTGAGCTGCTTCTTATCTGATAGCAAGCTGAAATGGCTCTTCTTCTATTACTCGCG	60		
Qy	61	TGGTTTCCAGTCTGCTCAGTCTTCTTCTCATGAGATTAAGGCTCTCCGTAGAACATCC	120		
Db	61	TGGTTTCCAGTCTGCTCAGTCTTCTTCTCATGAGATTAAGGCTCTCCGTAGAACATCC	120		
ORIGIN					
Qy	121	CAACTCTTGGAAATCTGACGGCCGGGAAATCGTTCGCGCATTCATAAACAATGTTTGA	180		
Db	121	CAACTCTTGGAAATCTGACGGCCGGGAAATCGTTCGCGCATTCATAAACAATGTTTGA	180		
Qy	181	CAAGCGTCGCATCTACTGATTCGTACAGAGCGGTGGGCAACTATCTATCCAACTGT	240		
Db	181	CAAGCGTCGCATCTACTGATTCGTACAGAGCGGTGGGCAACTATCTATCCAACTGT	240		
Qy	241	GGAGCGATGATTTTCATACAGTCTCTGATCTCAACGGCTTATGAGACCTGATTACCGG	300		
Db	241	GGAGCGATGATTTTCATACAGTCTCTGATCTCAACGGCTTATGAGACCTGATTACCGG	300		
Qy	301	AACGTGCTGACAGACTTATTGGGGAAGTAAAGATATATGTTCAATTTCAAGTCTGCTG	360		
Db	301	AACGTGCTGACAGACTTATTGGGGAAGTAAAGATATATGTTCAATTTCAAGTCTGCTG	360		
Qy	361	AAGATGGAGGCAATGATCTCTTCAACGACTTTTCTGCTCGATGACGTTGAACGTTTG	420		
Db	361	AAGATGGAGGCAATGATCTCTTCAACGACTTTTCTGCTCGATGACGTTGAACGTTTG	420		
Qy	421	GAAATCGAGGCAATTTCAAAAAGAGATAAAAACGGCACTCGATTTATGTAACGTTAT	480		
Db	421	GAAATCGAGGCAATTTCAAAAAGAGATAAAAACGGCACTCGATTTATGTAACGTTAT	480		
Qy	481	GGAAAGCAAAAAGGCAATTTGGATGTGGGAGGAGAGTGTGTGACTGACCTCAACTCAAC	540		
Db	481	GGAAAGCAAAAAGGCAATTTGGATGTGGGAGGAGAGTGTGTGACTGACCTCAACTCAAC	540		
Qy	541	CTTTGGGCTTCGAACTCTCCGACTACACGGATACACTGTGTCTTCAAGATGTTTGAAG	600		
Db	541	CTTTGGGCTTCGAACTCTCCGACTACACGGATACACTGTGTCTTCAAGATGTTTGAAG	600		
Qy	601	TTTTTAAAGACAAAATGSGCAATTTTCTCCACGCGCAATATTTCCAGGAGTCTCCCT	660		
Db	601	TTTTTAAAGACAAAATGSGCAATTTTCTCCACGCGCAATATTTCCAGGAGTCTCCCT	660		
Qy	661	TTAGAGGCTTCTCAATTTATTTCAGGGCTCCCTCGTCGCTTCCCGCGGAGAAAGTTA	720		
Db	661	TTAGAGGCTTCTCAATTTATTTCAGGGCTCCCTCGTCGCTTCCCGCGGAGAAAGTTA	720		
Qy	721	TGGATGAAGCTGAAACATTTCTCTACAAATATTTAAGAGAGCGCTGCAAAAGATTC	780		
Db	721	TGGATGAAGCTGAAACATTTCTCTACAAATATTTAAGAGAGCGCTGCAAAAGATTC	780		
Qy	781	CATCCAGTATATTTCTACTAGAGATACGGAGCTTCTGGAATATGTTGGCACACCAAT	840		
Db	781	CATCCAGTATATTTCTACTAGAGATACGGAGCTTCTGGAATATGTTGGCACACCAAT	840		
Qy	841	TGCCACGCTTGGAAAGCAAGAAATTTACATGGACGCTTTTGGACAGACACATAAAAT	900		
Db	841	TGCCACGCTTGGAAAGCAAGAAATTTACATGGACGCTTTTGGACAGACACATAAAAT	900		
Qy	901	ACGCGCGGAGAACTTTTGAACCTTGCAAAATTTGGAATTCATATATTTCACTCTTAC	960		
Db	901	ACGCGCGGAGAACTTTTGAACCTTGCAAAATTTGGAATTCATATATTTCACTCTTAC	960		
Qy	961	AAGAGAGAGAGTTTAAACATGTTTCCCGATGCTGGAAGACTCGGTTCTCTGAGATGA	1020		
Db	961	AAGAGAGAGAGTTTAAACATGTTTCCCGATGCTGGAAGACTCGGTTCTCTGAGATGA	1020		
Qy	1021	CTTCTCTGACATCGTCACGTGGAATTAATACGCTTTTGGCTTCTGCTTCTGCTGAGC	1080		
Db	1021	CTTCTCTGACATCGTCACGTGGAATTAATACGCTTTTGGCTTCTGCTTCTGCTGAGC	1080		
Qy	1081	CTCAACATTTCTGATTCAGACTCGGCTTTTACCAAGATGTCTCATCTTATCAOGTCT	1140		
Db	1081	CTCAACATTTCTGATTCAGACTCGGCTTTTACCAAGATGTCTCATCTTATCAOGTCT	1140		
Qy	1141	ACGACATGTACGACGCTCTTCGGCACAGTAGACGCTGGAACCTCTTACAGCGCAATTA	1200		
Db	1141	ACGACATGTACGACGCTCTTCGGCACAGTAGACGCTGGAACCTCTTACAGCGCAATTA	1200		



```
QY 713 GAAAGTTATGATGAAGCTGAACATCTCTACAAAATATTTAAGAGAGCCCTCAGAAA 772
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 707 GAAATTTATGATGAGGCTGAATCTTCTCTACAAAATATTTAAGAGAGCCCTCAGAAA 766
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 773 GATTCGGGATCCAGTATATCTTCACTAGAGATACGGACGTTCTTGGATATATGGTTGGCA 832
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 767 GATTCGGGCTCCAGT---CTTTCGGAGAGATCGGGACGTTTGGATATATGGTTGGCA 823
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 833 CACCAATTTGCCACGCTTGGAAAGCAAGAAATTAATGGAAGCTTTTGGACAGCAACATAA 892
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 824 CACATATTTTGGCGGATTTGGAAGCAAGAAATTAATCCTCAAGTCTTTTGGACAGCACTGA 883
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 893 AAATAAGAACGC-----CGCCGAGAAACTTTTAGAAGCTTCAAAATTTGGAATTCAA 943
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 884 GAACAGAGTCAATATGTAAGAGCAAAATCTTTTAGAAGCTCGCAAAATTTGGAATTCAA 943
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 944 TATATTTCACTCTTACAGAGAGAGAGTTAAACATATGTTTCCCGATGTTGGAAGAGCTC 1003
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 944 CATCTTTCAATCTTACAAAAGAGGAGTTAGAAAGTCTGGTCAGATGGTGAAGAATC 1003
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1004 GGGTTCTCTGAGATGACCTTCTGCGACATCGTCAATGGAATCTACGCTTTGGCTTC 1063
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1004 GGGTTTCTCGAGATGACCTTCTGCGACATCGTCAATGGAATCTACGCTTTGGCTTC 1063
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1064 CTGCATTCGGTTCGAGCTCAACATCTTGGATTCAGACTCGGCTTTACCAAGATCTCTCA 1123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1064 CTGCATTCGGTTCGAGCTCAACATCTTGGATTCAGACTCGGCTTTGCAAGAGCTGTC 1123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1124 TCTTATACGGTTCTTTCAGCATATGACGATCTTTCGGACAGTAGAGCTGGAATC 1183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1124 TCTTATACGGTTCTTTCAGCATATGACGATCTTTCGGACAGTAGAGCTGGAATC 1183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1184 CTTTCAGCGCAATTAAGAGATGGATCGGTCGCGATGGAATGCCCTTCCAGATATAT 1243
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1184 CTTTCAGCGCAATTAAGAGATGGATCGGTCGCGATGGAATGCCCTTCCAGATATAT 1243
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1244 GAAAGAGTGTACATGATGTTTATCACACGCTAAATGAATGCTCGAGTGGCAGAGAA 1303
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1244 GAAAGAGTGTACATGCGGTTTACGACCGCTAAATGAATGCTCGAGTGGCAGAGAA 1303
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1304 GGCTCAAGCGCCGAGACAGCTCAATATSCAAGACAGAGCTTGGAGGCGTGTGTTGATTC 1363
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1304 GGCTCAAGCGCCGAGATACGCTCAATATGCTCGGAAGCTTGGAGGCTTATATGATTC 1363
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1364 GTATATGAGAGCAAGCAAGTGTATCGCCACTGGTTATCTGCCACGTTTGGAGGATCTT 1423
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1364 GTATATGAGAGCAAGGTTGATCGCCACTGGTTATCTGCCACGTTTGGAGGATCTT 1423
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1424 GGAGAACGGGAAAGTTAGCTCTGCTCATCGCCCATGCGCACTGCAACCCATCTTGACGTT 1483
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1424 CGAATGGGAAAGTTAGCTGTGTCATCGCATATCCGATGCAACCCATCTTGACAAAT 1483
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1484 GGAATCCCTTTCTGATCATCTCTCAAGGAAGTTGACTTCCCATCGAAGCTCAATGA 1543
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1484 GGAATCCCTTTCTGATCATCTCTCAAGGAAGTTGACTTCCCATCGAAGCTCAATGA 1543
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1544 CTTTCATATGATCATCTCTCAATTAAGAGGTATACACGTTGCTCAAGGCAAGCAGGCG 1603
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1544 CTTTGGCATGTGCCATCTTCGATTAACGAGGTGATACGCGTGTCTACAGGCGGACAGGCG 1603
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1604 CCGTGGAGAGAAAGCTTCGTATATCATGTTTATATGAAGACAACTCTCGAATTAACGGA 1663
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1604 TCGTGGAGAGAAAGCTTCGTATATCATGTTTATATGAAGACAACTCTCGAATTAACGGA 1663
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1664 AGAAGATGCTCTGAATCATATCAACTTCAATGATCAGGAGCAATCAGAGAAATTAATTTG 1723
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1664 GGAAGATGCTCTGATCATATCAAGCCATGATCAGTCACTANTCAAGAGATTAATTTG 1723
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1724 GGAGCTTCTAAGCCAGACACAGTGTTCCTCATCACTTCCAGAAACACGCAATTTGACAT 1783
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1724 GGAACTTCTCAAAACAGACATCAATGTTTCCATCTCGGCGAAGAAACATGCTTTTGACAT 1783
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1784 AAGCAGAGTTTGGCATCAGGTTTACAGATACCGAGATGGCTTACAGCTTTGCGCAAGTTGA 1843
```

```
Db 1784 CGCCAGAGCTTTCCATTACGGCTACAAATACCGACGCGCTACAGCGTTGCCAAGCTTGA 1843
QY 1844 AACAAAGAGTTTGGTATGAGAACCGTCATTAAGAACTGTGCCTTTGTACACACTTCAA 1903
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1844 AACGAAGAGTTTGGTATGAGAACCGCTCTTGAATCTGTGCCTTTGTAGCAACAGCTCAA 1903
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1904 ATCTACAATATTAACCTGAGGATGCCCTATGGGTGTATATAGGGCACACAAAAATAAATAT 1963
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1904 ATCTATGCCCTATGCTATGTCGGGTTAAATATATGTGGAAGGTAGCGCTTGAGTGTAGA 1963
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1964 GGTCTGTGTAGTAAGCTGTAAATTTATGAAAAAATAAAAAAATAAAAAA 2013
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1964 GATAAGTTTGTATTAATTTAAAGTTGTAATTTAAAAAATAAAAAAATAAAAAA 2013
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
AR222097 Locus AR222097 2018 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 3 from patent US 6429014.
ACCESSION AR222097
VERSION AR222097.1 GI:23329471
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2018)
AUTHORS Steele,C.L., Bohlmann,J. and Croteau,R.B.
TITLE Monoterpene synthases from grand fir (Abies grandis)
JOURNAL Patent: US 6429014-A 3 06-AUG-2002;
FEATURES
Location/Qualifiers
source
1..2018
/organism="unknown"
/mol_type="genomic DNA"

Query Match 64.9%; Score 1306.8; DB 6; Length 2018;
Best Local Similarity 81.0%; Pred. No. 0;
Matches 1595; Conservative 0; Mismatches 342; Indels 33; Gaps 5;

QY 68 CAGGTCTGTCCTCAGTTCTTCTCATGAGATTAAAGGCTCTCGTAGAACAAATCCCACTCT 127
Db 53 CAATCGTTGATCAGTTCTACCCATGAGCTTAAGGCTCTCTAGAACAAATTCAGCTCT 112
QY 128 TGAATCTGACGGCGGGAATCCGTCGCAATTCATAAACAATGTGTTTCAACAAGCGT 187
Db 113 AGGAATGAGTAGGCGAGGAAATCTATCACTCTTCCATCAGCATGAGCTCTACCAAGT 172
QY 188 CGCATCTACTGATCTGTACAGACGCGTGGGCACTATCATTCACACCTGTGGGACGA 247
Db 173 TGTAAACCGATGATGGTGTACGAAGACGATGGGCGATTTCATTCACACCTCTGGGACGA 232
QY 248 TGATTTTCATACAGTCTCTGATCTCAACGCTTATGGAGCACTGATTTACCGGAAAGCTGC 307
Db 233 TGATGTCATACAGTCT---TTACCAAGCGCTTATAGGAAAAATCGTACCTGGAGCGTGC 289
QY 308 TGACAGACTTATTTGGGGAAGTAAAGGATATAATGTTCAAATTTCAAGTCGCTGGAGATGG 367
Db 290 TGAGAACTGATCGGGAAGTAAAG---AACATGTTCAATTCGATGTCAATTAGAGATGG 346
QY 368 AG-----GCAATGATCTCTTTCAACGACTTTTGTGCTGGTTCGATGAGCTTGA 412
Db 347 AGATTTAATGAGTCCGCTCAATGATCTCAATTCACGCTTTGGATTTCGACAGCTTTGA 406
QY 413 ACCTTTTCGGGAATCGACAGGCAATTTCAAAAAGAGATAAAAACGCACTCCGATTAATGTTAA 472
Db 407 ACCTTTTCGGGAATCGATAGACATTTCAAAAGATAGATAAAATTCGGCGCTTGATTAATGTTTA 466
QY 473 CAGTTATTGAAACGAAAAAGGCAATTTGGATGTGGGAGGAGAGTGTGCTGATGACCTCAA 532
Db 467 CAGTTATTGGGCGAAATGGCATCGATCGATCGGAGGAGGAGTGTGTTGTTACTGATCTGAA 526
QY 533 CTCAACGCGCTTGGGCTTCGAACTCTCCGACTACCGGATACACTGTGCTTTCAGATGT 592
```

527 CTCACCTGGCTTGGGCTTCGACCCCTACGACTACAGGATACCCGGTGTCTTCAGATG 586  
 593 TTTGAAAGTTTAAAGACAAATAATGGGCAATTTTCTCCACTGCGCAATATTCAGATGA 552  
 587 TTTCAAAGCTTTCAAAGGCCAAATGGGAGTTTCTCTGCTCTGAAATATTCAGACAGA 646  
 653 GGGAGAGATTAGAGGGGTTCTCAATTTATTCAGGGGCTCCCTGCTGCTCTTCCGGGCA 712  
 647 TGAAGAGATCAGAGGGGTTCTGAAATTTATTCGGGGCTCCCTCATTTGCGCTTTCCAGGGGA 706  
 713 GAAAGTTATGGATGAAGCTGAAACATTTCTCAAAAATATTTAAAGAGAGCCCTGCAAAA 772  
 707 GAAATATATGGATGAGCTGAATTTCTCTACCAATATTTAAAGAGAGCCCTGCAAAA 766  
 773 GATTCGGGATCCAGATATCTTTCACTAGAGATACGGGAGCTTCTGGAATATGGTTGGCA 832  
 767 GATTCGGGCTCCAGT---CTTTCGGGAGAGATCGGGGAGCTTTTGGAAATATGGTTGGCA 823  
 833 CACCAATTTGCCAGCTTTCGAGCAGAGGATTCATGAGGCTCTTTGGACAGCACACTAA 892  
 824 CACATATTTCCGCGGATTTGAAGCAAGGAATTCATCCAAAGTCTTTTGGACAGGACACTGA 893  
 893 AAATAAGAACGC-----CGCCGAGAACTTTTAGAACTTGCAAAATTTGGAATTCAA 943  
 884 GAACAGAGATCATATGTGAGAGCAAAAACCTTTTAGAACTCGCAAAATTTGAGTTCAA 943  
 944 TATATTCACTCTTCAAGAGAGAGAGTTAAACATGTTTCCGATGGTGGGAAAGACTC 1003  
 944 CATCTTTCAATCTTCAAAAGAGGGAGTTAGAAAGTCTGGTCAGATGGTGGGAAAGATC 1003  
 1004 GGGTCTCTCGAGATGACCTTCTGCGACATCTGACGATGGAATCTAGCTTTGGCTTC 1063  
 1004 GGGTCTCTCGAGATGACCTTCTGCGACATCTGACGATGGAATCTAGCTTTGGCTTC 1063  
 1064 CTGCATTTGCTTCGAGCTCAACATTTCTGATTCAGACTCGGCTTTACCAAGATGTCTCA 1123  
 1064 CTGCATTTGCTTCGAGCTCAACATTTCTGATTCAGACTCGGCTTTGCCAAGACGTGCTCA 1123  
 1124 TCTTATCAGGTTTCTTGAAGATGACGCTTCTGCGACATCTGACGATGGAATCTAGCTTTG 1183  
 1124 TCTTATCAGGTTTCTTGAAGATGACGCTTCTGCGACATCTGACGATGGAATCTAGCTTTG 1183  
 1184 CTTCAGAGCACAATTAAGAGATGGGATCGTCCGCGATGGAATGCTTCCAGAAATAT 1243  
 1184 CTTCAGAGCACAATTAAGAGATGGGATCGTCCGCGATGGAATGCTTCCAGAAATAT 1243  
 1244 GAAAGGAGTGTACATGATGTTTATCACACCGTAAATGAAATGGCTCGAGTGGCAGAGAA 1303  
 1244 GAAAGGAGTGTACATGATGTTTATCACACCGTAAATGAAATGGCTCGAGGCGAGAGGA 1303  
 1304 GGGTCAAGGCGGAGACACGCTCAACTATGCAAGACAGGCTTGGAGGCGTGTGTTGATTC 1363  
 1304 GGGTCAAGGCGGAGATACGCTCAACTATGCTCGGGAAGCTTGGAGGCTTATATGATTC 1363  
 1364 GTATATCAGAGAGCAAGTGGATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1423  
 1364 GTATATGCAAGAGCAAGTGGATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1423  
 1424 GGAAGAGCGGAAAGTTAGCTCTGCTCATCGCCCATGCGCACTGCAACCCATTTCTGAGCTT 1483  
 1424 CGAAGATGGGAAAGTTAGCTCTGCTCATCGCCCATGCGCACTGCAACCCATTTCTGACAT 1483  
 1484 GGCATCCCTTCTGATCATCTTCAAGAGGATGATCTCCATTCGAGCTCAATGA 1543  
 1484 GGCATCCCTTCTGATCATCTTCAAGAGGATGATCTCCATTCGAGCTCAATGA 1543  
 1544 CTTGATATGATCATCTTCAAGAGGATGATCTCCATTCGAGCTCAATGA 1603  
 1544 CTTGATATGCTTCTGATCATCTTCAAGAGGATGATCTCCATTCGAGCTCAATGA 1603  
 1604 CCGTGGAGAGAGCTTCTGCTTATCATGTTTATGAAGACATCTGATTAACGGA 1663

1604 TCGTGAGAGAGAGCTTCTCTATATCATGTTTATGAAGAGCAATCTCTGGAGTATCAGA 1663  
 1664 AGAGATGCTCTGAATCATATCATCACTTCAATGATCAGGAGCGCAATCAGAGAAATTAATTTG 1723  
 1664 GGAAGATGCTCTCGATCATATCAACGCCATGATCAGTGACGTAAATCAAGAGTAAATTTG 1723  
 1724 GGAGCTTTTAAAGCCAGACACAGAGTGTTCCTCATCTTCCAAAGAAACACGCAATTTGACAT 1783  
 1724 GGAATCTTCAACCCAGACATCAATGTTTCCCATCTCGGGAAGAAACATGCTTTTGACAT 1783  
 1784 AGAGAGATTTGGCATCAGGTTTACAGATACCGAGATGCTTACAGCTTTGCCACGTTGA 1843  
 1784 CGCCAGAGCTTTCCATTAACGCTTCAAAATACCGAGAGCGCTACAGGCTTGCACACGTTGA 1843  
 1844 AACAAAGCTTTGGTATGAGAACCGCTCAATGAACCTGTGCTTCTTAAACACACATTTCAA 1903  
 1844 AACGAAGATTTGGTACGAGAACCTCTCTGAACTGTGCTTGTGCTTTGTAGCAACGCTCAA 1903  
 1904 ATCTACAATATTAACCTGAGGATGCCCTATGGGTGTATATAGGCGACACAAAAATAATAT 1963  
 1904 ATCTATGCCCTATGCTATGTCGGTTTAAATATATGCTGAAGGTAGCCGTTGGATGAGA 1963  
 1964 GGTGTGTAGTAAAGCTGTATTTATGAAAAAATAAATAAATAAATAAATAAATAAATAA 2013  
 1964 GGATAAGTTTGTATTAATTAATAAAGTTGTAATTTTAAAAAATAAATAAATAAATAAATAA 2013

RESULT 5  
 LOCUS AR240696 2018 bp DNA linear PAT 20-DEC-2002  
 DEFINITION Sequence 19 from patent US 6468772.  
 ACCESSION AR240696  
 VERSION AR240696.1 GI:27285845  
 KEYWORDS Unknwn.  
 SOURCE Unknwn.  
 ORGANISM Unknwn.  
 REFERENCE 1 (bases 1 to 2018)  
 AUTHORS Chappell, J., Manna, K.R., Noel, J.P. and Starks, C.M.  
 TITLE Methods of making modified polypeptides  
 JOURNAL Patent: US 6468772-A 19 22-OCT-2002;  
 FEATURES Location/Qualifiers  
 source  
 1..2018  
 /organism="unknwn"  
 /mol\_type="genomic DNA"

ORIGIN  
 Query Match 64.9%; Score 1306.8; DB 6; Length 2018;  
 Best Local Similarity 81.0%; Pred. No. 0;  
 Matches 1595; Conservative 0; Mismatches 342; Indels 33; Gaps 5;

QY 68 CAGGTGCTGCTCAGTTCTCTCATGAGATTAAGGCTCTCCGTAGAACAAATCCCACTCT 127  
 Db 53 CAAATCGTGTGATCAGTTCTACCCATGAGCTTAAGGCTCTCTTAGAACAAATCCAGCTCT 112  
 QY 128 TGAATCTGAGGCGCGGGAAATCCGTGCGCAATCCATAAAATGTTTGAACGCT 187  
 Db 113 AGGAATGAGTGTAGCGGAGGAAATCTATCTCTCCATCAGCATGAGCTCTACCACT 172  
 QY 188 CGCATCTACTGATTCGTACAGAGCGCTGGGCAACTATCATTCCAACCTGTGGGACGA 247  
 Db 173 TGTAAACCGATGATGGTGTACGAAGACGATGGGCGAATTCCTATCCAACTCTGGGACGA 232  
 QY 248 TGATTTTCATACAGTCTCTGATCTCAACGCTTATGGAGCACTTGAATACCGGAAACGTCG 307  
 Db 233 TGAATGTCATACAGTCT---TTACCAAGGCTTATGAGGAAATCGTACCTGGAGCGTGC 289  
 QY 308 TGACAGACTTATTTGGGAGTAAGGATATATGTTCAATTTCAAGTCGCTGGAAGATGG 367  
 Db 290 TGAGAAACTGATCGGAGAGTAAAG---AAACATGTTCAATTCGATGTCATTAGAAGATGG 346  
 QY 368 AG-----GCAATGATCTCTTCAACGACTTTTGGTGGTGCATGATGAGTTGA 412

```
Db 347 AGAGTTAATGAGTCCGCTCAATGATCTCATTTCAACGCCTTTTGGATTGTCGACAGCCTTGA 406
QY 413 AGTTTGGGAATCGACAGGCATTTCAAAAAGAGATAAAAAACGGCACTCGATTATGTTAA 472
Db 407 AGTTTGGGGAATCCATAGACATTTCAAAAGATGAGATAAAATCGGCGCTGATTATGTTTA 466
QY 473 CAGTTATTGGAACGAAAAAGGATTTGGATGTGGGAGGAGAGTGTGTGACTGACCTCAA 532
Db 467 CAGTTATTGGGCGGAAAAATGGCATCGGATGCGGGAGGGAGAGTGTGTGTACTGATCTGAA 526
QY 533 CTCACCGGCTTTGGGCTTCGAACTCTCCGACTACACGGATACACTGTGTCTTTCAGATGT 592
Db 527 CTCAACTGCGTTGGGCTTCGAACCTCAGACTACCGGATACCGGCTCTTCAGATGT 586
QY 593 TTTGAACGTTTAAAGACAAAAATTTGGGCAATTTTCCTCCAATGCGCAATATTCAGATAGA 652
Db 587 TTTCAAAAGCTTTCAAAGGCAAAATTTGGGCAATTTTCCTGCTCTGAAAAATATTCAGACAGA 646
QY 653 GGGAGAGATTAGAGCGCTTCTCAATTTATTCAGGCGCTCCCTCGTCGCTTCCCGGCGA 712
Db 647 TGAAGAGATCAGAGCGCTTCTCAATTTATTCGCGGCTCCCTCATTTGCTTTCCAGGGGA 706
QY 713 GAAAGTTATGAGATGAGCTGAAACATTTCTCAAAAATATTTAAGAGAGCGCTGCAAAA 772
Db 707 GAAAAATATGAGATGAGCTGAAATCTTCTTACCAAAATATTTAAAGAGCGCTGCAAAA 766
QY 773 GATTCGCGATCCAGTATACATTTTACTAGAGATACGGAGCGTCTTGGAAATATGTTGGCA 832
Db 767 GATTCGCGCTCCAGT---CTTTCGCGAGAGATCGGGAGCGTTTGGAAATATGTTGGCA 823
QY 833 CACCAATTTGGCACGCTTGGAGCAAGGAATTTACATGACGCTCTTGGACAGCACTAA 892
Db 824 CACATATTTGCGGATTTGGAAGCAAGGAATTTACATCCAGTCTTTGGACAGGACACTGA 883
QY 893 AAATAAGAACGC-----CGCGAGAAACTTTTGAACCTTGCAAAATTTGGAATTTCAA 943
Db 884 GAACAGAAAGTCAATGATGGAAGAGCAAAAAACTTTTAGAATCTCGCAAAATTTGGAGTTCAA 943
QY 944 TATATTTCACTCCTTACAAGAGAGAGATTAAACATGTTTCCGATGGTGGAAAGACTC 1003
Db 944 CATCTTTCAATCCTTTACAAAAAGAGGAGTTAGAAAGTCTGGTCAGATGGTGGAAAGATC 1003
QY 1004 GGGTTCTCCTCAGATGACCTTCTGTCGACATCGTCACGTGGAATACACTACGCTTTGGCTTC 1063
Db 1004 GGGTTTCTCAGATGACCTTCTGCCGACATCGTCACGTGGAATACACTACCTTTGGCTTC 1063
QY 1064 CTGCATTTGCGTTCGAGCTCAACATTTTGGATTGAGCTCGGCTTTTACCAAGATGTCTCA 1123
Db 1064 CTGCATTTGCGTTCGAGCTCAACATTTTGGATTGAGCTCGGCTTTTGCAGAGCTGTCA 1123
QY 1124 TCTTATCAGCTTCTTACGACATGTAGGAGCTTTCGGACAGTAGACGCTTGGAACT 1183
Db 1124 TCATTATCAGCTTCTTACGATATGTAGCAGACCTTCGGCACAGTAGACGCTTGGAACT 1183
QY 1184 CTTTACAGCGACAAATTAAGAGATGGGATCCGCTCGCGATGGAATGCTTCCAGAAATATAT 1243
Db 1184 CTTTACAGCGACAAATTAAGAGATGGGATCCGCTCGATAGATTGCTTCCAGAAATATAT 1243
QY 1244 GAAAGAGTGTACATGATGTTTATCACACGTAATGAAATGAAATGCTCGAGTGGCAGAGAA 1303
Db 1244 GAAAGAGTGTACATGATGTTTATCACACGTAATGAAATGAAATGCTCGAGAGGACAGAGAA 1303
QY 1304 GGCTCAAGGCGGACAGACGCTCACTATGCAAGACAGGCTTGGAGGGGCTTTTGTATTC 1363
Db 1304 GGCTCAAGGCGGACAGATACGCTCAATATGCTCGGGAAGCTTGGAGGGCTTATATTGATTC 1363
QY 1364 GTATATCAGGAACAAAGTGGATGCGCAGCTGTTATCTGCCACGTTTGGAGGATGACTT 1423
Db 1364 GTATATCAGGAACAAAGTGGATGCGCAGCTGTTATCTGCCCTCTTCTTATGATGACTA 1423
QY 1424 GGAGAACGGAAAGTTAGCTCTGCTCATCGCCCATGCGCACTGCAACCCCAATTTCTGAGTT 1483
Db 1424 CGAATGGGAAGTTAGCTGTGCTCATCGCATATCGCATTTGCAACCCCAATTTCTGACAAAT 1483
```

```
QY 1484 GGACATCCCTTTCTGATCAGATCCTCAAGGAAGTTGACTTCCCATCGAAGCTCAATGA 1543
Db 1484 GGACATCCCTTTCTGATCAGATCCTCAAGGAAGTTGACTTCCCATCGAAGCTCAATGA 1543
QY 1544 CTTGATATGATCATCTTCTGATTAAGAGTGTATACAGGTCTCAAGGCGAGACAGGCG 1603
Db 1544 CTTGATATGATCATCTTCTGATTAAGAGTGTATACAGGTCTCAAGGCGAGACAGGCG 1603
QY 1604 CGTTCGAGAGAGCTTCTGATCATGTTATATGATTAAGAGCAATCTCTGGAGTATCAGA 1663
Db 1604 TCGTGGAGAGAGCTTCTCTATATCATGTTATATGATTAAGAGCAATCTCTGGAGTATCAGA 1663
QY 1664 AGAAGATCTCTGATCATATCAACTTTCATGATCAGGAGCGCAATCAGAGAAATTAATTTG 1723
Db 1664 GGAAGATCTCTGATCATATCAAGCCCATGATCAGTACGCTTAATCAAGAGTTTAAATTTG 1723
QY 1724 GGAGCTTCTAAAGCCAGACACAGTGTTCCTATCACTTCCAAAGAAACCGCATTTGACAT 1783
Db 1724 GGAATCTCTCAACCCAGACATCAATGTTCCCATCTCGCGAAGAAACATGCTTTTGTGACAT 1783
QY 1784 AAGCAGAGTTTGGCATCACGTTACAGATACCGAGATCGCTTGTAAACACACTTCAA 1843
Db 1784 CGCCAGAGCTTTCATATACGCTCAAAATACCGAGACGCTACAGCTTGGCAACGTTGA 1843
QY 1844 AACAAAGAGTTTGGTGTAGAGAACCGTCAATCAACCTTGTGCTTTGTAAACACACTTCAA 1903
Db 1844 AACGAAGAGTTTGGTTCACGAGAACCTTCTTGAATCTGTGCTTTGTAGCAACAGCTCAA 1903
QY 1904 ATCTACAAATATTAAGTGTAGGATGCGCTTATGAGGTGTATATAGGGCACACAAAATTAATAT 1963
Db 1904 ATCTATGCCCTATGCTATGTCGGGTAAATATATATGTGGAAGGTAGCGCTTGGATGTAGA 1963
QY 1964 GGTGTGTGTAGTAAAGCTGTAAATTTATGAAAAAATAAATAAATAAATAAATAAATAA 2013
Db 1964 GGATAAGTTTGTATTAATTAATTAAGTGTAAATTTAATAAATAAATAAATAAATAAATAA 2013

RESULT 6
AR266971
LOCUS AR266971 2018 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 19 from patent US 6495354.
ACCESSION AR266971
VERSION AR266971.1 GI:29696426
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2018)
AUTHORS Chappell,J., Manna,K.R., Noel,J.P. and Starks,C.M.
TITLE Syntheses
JOURNAL Patent: US 6495354-A 19 17-DEC-2002;
FEATURES
source 1..2018
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 64.98; Score 1306.8; DB 6; Length 2018;
Best Local Similarity 81.08; Pred.No. 0;
Matches 1595; Conservative 0; Mismatches 342; Indels 33; Gaps 5;

QY 68 CAGTCTCGTCTCAGTTCTTCTCATGAGATTAAGGCTCTCCGTAGAGAAATCCCAACTCT 127
Db 53 CAAATCGTTGATCAGTCTTACCCATGAGCTTAAGGCTCTCTCTAGAACAAATCCAGCTCT 112
QY 128 TGAATCTGAGCGCGGGGAAATCCGTCGCGAATTCATATAACATGTTGTTGACAACGCT 187
Db 113 AGGAATGAGTAGGAGGAGGAAATCTATCACTCCTTCCATCAGCATGAGCTCTACCCCGT 172
QY 188 CGCATCTACTGATTCGTACAGAGACGCGTGGCGCACTATCATCCACCTGTGGGACGA 247
Db 173 TGTAAACGATGATGTTGTACGAAGACGATGGGCGGATTTCCATTCCAACCTCTGGGACGA 232
```





JOURNAL Patent: US 6569656-A 19 27-MAY-2003;  
FEATURES Location/Qualifiers  
1..2018  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 64.9%; Score 1306.8; DB 6; Length 2018;  
Best Local Similarity 81.0%; Pred. No. 0;  
Matches 1595; Conservative 0; Mismatches 342; Indels 33; Gaps 5;  
QY 68 CAGGTCGTCCTCAGTCTCTCTCATGAGATTAAAGCTCTCCCTAGAACAAATCCCAACTCT 127  
DB 53 CAAATCGTTGATCAGTCTTACCCATGAGCTTAAAGCTCTCTCTAGACAAATCCAGCTCT 112  
QY 128 TGGAAATCTGACGCCGGGAAATCCGTCGCGATTCCATTAACATGTTGTCAGAGCGT 187  
DB 113 AGGAATGAGTAGCGGAGGAAATCTATCACTCTTCCATCAGCATGAGCTCTACCAACGT 172  
QY 188 CCGATCTACTGATTTCTGTACAGAGCGCTGGGCACTATCATTCCAACCTGTGGACGA 247  
DB 173 TGTAAACCGATGATGTTGACGAGACGCAATGGCGGATTTCCATTCCAACTCTGGACGA 232  
QY 248 TGATTTTATACAGTCTCTGATCTCAACGCTTTATGAGCACCTGATTACCGGAAACGTCG 307  
DB 233 TGATGTCATACAGTCT--TTACCAACGCTTATGAGGAAATCGTACCTGGAGCGTGC 289  
QY 308 TCACAGACTTTATGGGGAAGTAAAGGATTAATGTTCAATTTCAAGTCTGAGAGATGG 367  
DB 290 TGAGAACTGATCGGGGAAGTAAAG---AACATGTTCAATTCGATGTCATTAGAAATGG 346  
QY 368 AG-----GCAATGATCTCTTCAACGACTTTTCTGCTGATGACGTTGA 412  
DB 347 AGAGTTTAATGATTCGCTCAATGATCTCATTTCAACGCTTTGATTTGTCGACGCTTGA 406  
QY 413 ACCTTTTGGGAATCGACAGGCAATTTCAAAAAGAGATAAAACGGCACTCGATTATGTTAA 472  
DB 407 ACCTTTGGGATCCATAGACATTTCAAGATGAGATAAAATCGCGCTTGATTATGTTA 466  
QY 473 CAGTTATTCGACGAAAGGCAATTTGATGTTGGAGGAGAGTGTGTGACTGACCTCAA 532  
DB 467 CAGTTATTCGGGCGAAATGGCATCGGATCGGGGAGGAGTGTGTTACTGATCTGAA 526  
QY 533 CTCACCGCTTGGGCTTCGAACTCTCGACTACAGGATACACTGTGCTTCAGATGT 592  
DB 527 CTCACCTGCTTGGGCTTCGAACTCTCACTACAGGATACCGGTGCTTCAGATGT 586  
QY 593 TTTTGAACGTTTTTAAAGACAAAATGGCAATTTTCTCCACTGCCAATATTCAGATAGA 652  
DB 587 TTTTCAAGCTTTCAAGGCCAAAATGGCAGTTTTCTGCTCTGAAATATTCAGACAGA 646  
QY 653 GGGAGAGATTAGAGGCGTCTCAATTTATTCAGGGCTCTCCCTGCTCGCTTTCCCGCGA 712  
DB 647 TGAAGAGATCAGAGGCGTCTGAAATTTATTCGGGCTCTCCCTCATTCCTTTCCAGGGGA 706  
QY 713 GAAAGTTATGATGAGCTGAAATCTCTTCAAAAATTTTAAAGAAAGCCCTGCAAAA 772  
DB 707 GAAATTTATGATGAGCTGAAATCTCTTCAAAAATTTTAAAGAAAGCCCTGCAAAA 766  
QY 773 GATTCCGGCATCCAGTATATCTTCACTAGAGATACGGGAGCTTCTGGAATATGGTTGGCA 832  
DB 767 GATTCCGGCTCCAGT---CTTTCGCGAGAGATCGGGGAGCTTTTGGAAATATGGTTGGCA 823  
QY 833 CACCAATTTGCCAGCTTGAAGCAAGAAATTTACATGAGCTCTTTTGGACAGCACTAA 892  
DB 824 CACATATTTGCCGATTTGGAGCAAGAAATTTACATCAAGTCTTTTGGACAGGACACTGA 883  
QY 893 AAATAAGAACGC-----CGCCGAGAACTTTTGAACCTTCCAAAATTTGGAATTCAA 943  
DB 884 GAAACAGAGTCATATGTGAAGACGAAATTTTGAACCTCGCAAAATTTGGAGTTCAA 943  
QY 944 TATATTTCACTCTTTACAGAGAGAGTTTAAACATGTTTCCCGATGGTGGAAAGACTC 1003

DB 944 CATCTTTCAATCCTTTACAAAAGAGGGAGTTAGAAAGTCTGGTCAGATGGTGGAAAGATC 1003  
QY 1004 GGGTTCTCTGAGATGACCTTCTGTGCGACATCGTCACGTTGGAAATACTACGTTTGGCTTC 1063  
DB 1004 GGGTTTCTGAGATGACCTTCTGCGACATCGTCACGTTGGAAATACTACACTTGGCTTC 1063  
QY 1064 CTGCATTGGTTTCGAGCCTCAACATTTCTGGATTGAGCTCGGCTTTGCCAAGACGTTGCA 1123  
DB 1124 TCTTATCAGGTTCTTTCGACATCTAGCAGCTCTTTCGSCACAGTAGACGAGCTGAACT 1183  
QY 1124 TCTTATCAGGTTCTTTCGACATCTAGCAGCTCTTTCGSCACAGTAGACGAGCTGAACT 1183  
DB 1184 CTTTCACAGCGCAATTTAAGAGATGGATCCGTCGCGATGGAATGCCCTTCCAGAAATATAT 1243  
QY 1184 CTTTCACAGCGCAATTTAAGAGATGGATCCGTCGCGATGGAATGCCCTTCCAGAAATATAT 1243  
DB 1244 GAAAGAGTGTATCATGATGGTTTATCAACCGTAAATGAAATGGCTCGAGAGGCGAGAA 1303  
QY 1244 GAAAGAGTGTATCATGATGGTTTATCAACCGTAAATGAAATGGCTCGAGAGGCGAGAA 1303  
DB 1304 GGCTCAAGCCGAGACAGCTCAACTATGCAAGACAGGCTTGGAGGCGCTGTTTGAATTC 1363  
QY 1304 GGCTCAAGCCGAGATACCGTCAATATGCTCGGAGCTTGGAGGCTTATATTCATTC 1363  
DB 1364 GTATATGCAAGGAAAGTGGATCGCCACTGTTATTCGCCACAGCTTGGAGAGTACTT 1423  
QY 1364 GTATATGCAAGGAAAGTGGATCGCCACTGTTATTCGCCACAGCTTGGAGAGTACTT 1423  
DB 1424 GAGAAACGGGAAAGTTAGCTCTGCTCATCGCCCATCGGCACTGCAACCCATCTGACGTT 1483  
QY 1424 CGAAGATGGGAAAGTTAGCTGCTGCTATCGCATATCCGATTTGCAACCCATCTGCAAT 1483  
DB 1484 GGCATCCCTTCTTCGATCATCTCAAGGAAGTTGACTTCCCATCGAAGCTTAAACGA 1543  
QY 1484 GGCATCCCTTCTTCGATCATCTCAAGGAAGTTGACTTCCCATCGAAGCTTAAACGA 1543  
DB 1544 CTTGATATGATCATCTTCGATTAAGAGTGTATACACGCTGTATCAAGGACAGAGGC 1603  
QY 1544 CTTGATATGATCATCTTCGATTAAGAGTGTATACACGCTGTATCAAGGACAGAGGC 1603  
DB 1604 CCGTGAGAGAAAGCTTCTGCTATATCATGTTATGAAAGCAATCTCGATTAAACGA 1663  
QY 1604 TCGTGAGAGAAAGCTTCTGCTATATCATGTTATGAAAGCAATCTCGATTAAACGA 1663  
DB 1664 AGAAGATGCTCTGAATCATATCACTTCATGATCAGGAGCGCAATCAGAAATTAATTTG 1723  
QY 1664 GGAAGATGCTCTGATCATATCAACGCCATGATCAGTACGTAATCAAGGATTAATTTG 1723  
DB 1724 GGAGCTTTTAAAGCCAGACAAACAGTGTTCCTCATCTTCCATCATCTTCCATCATCTT 1783  
QY 1724 GGAATTTCTCAAAACCCAGACATCAATGTTCCCATCTCGGCGAAGAAACATCTTTTGCAT 1783  
DB 1784 AAGCAGAGTTTGGATCAGATACCGAGATGGCTACAGCTTTTCCCAACGTTGA 1843  
QY 1784 CGCCAGAGCTTTCATTAACGCTTACAAATACCGAGAGCGCTACAGCTTTCGCAACGTTGA 1843  
DB 1844 AACCAAGAGTTTGGTGAATGAGAAACCGCTCATTTGAACCTGCTTCTTTGTAACAACTTCAA 1903  
QY 1844 AACCAAGAGTTTGGTGAATGAGAAACCGCTCATTTGAACCTGCTTCTTTGTAACAACTTCAA 1903  
DB 1904 ATCTCAATATTAATCTGAGGATGCCCTATCGGCTTATATGAGGAGGAGTCCGTTGATGTAGA 1963  
QY 1904 ATCTATGCCCTATGCTATGTCGGGTTAAATATATATGTTGAGAGGAGTCCGTTGATGTAGA 1963  
DB 1964 GGTGTGTTAGTAAAGCTGTAATTTATGAAAAAATAAATAAATAAATAAATAAATAAATAA 2013  
QY 1964 GGTGTGTTAGTAAAGCTGTAATTTATGAAAAAATAAATAAATAAATAAATAAATAAATAA 2013  
DB 1964 GGATAAGTTTGTATTAATTTTAAATAGTTGTAATTTTAAATAAATAAATAAATAAATAA 2013

LOCUS	AR429869	2018 bp	DNA	linear	PAT 18-DEC-2003
DEFINITION	Sequence 19 from patent US 6645762.				
ACCESSION	AR429869				
VERSION	AR429869.1	GI:40190267			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 2018)				
AUTHORS	Chappell, J., Manna, K.R., Noel, J.P. and Starks, C.M.				
TITLE	Synthases				
JOURNAL	Patent: US 6645762-A 19 11-NOV-2003;				
FEATURES	Location/Qualifiers				
source	1..2018				
	/organism="unknown"				
	/mol_type="genomic DNA"				
ORIGIN					
	Query Match	64.9%;	Score 1306.8;	DB 6;	Length 2018;
	Best Local Similarity	81.0%;	Pred. No. 0;		
	Matches 1595;	Conservative 0;	Mismatches 342;	Indels 33;	Gaps 5;
QY	68	CAGTCGTCGCTCAGTCTCTCTCATGAGATTAAGGCTCTCGGTAGAACAAATCCCAACTCT	127		
DB	53	CAAAATCGTTGATCAGTCTCTACCCATGAGCTTAAGGCTCTCTCTAGAACAAATCCAGCTCT	112		
QY	128	TGGAATCTGACGGCGGGGAAATCCGTCGCGCAATTCATAAACAATGTTTGGACAAGCGT	187		
DB	113	AGGAATGAGTAGGCGAGGAAATCTATCACTCTTCCATCAGCATGAGCTCTACCAACCGT	172		
QY	188	CGCATCTACTGATCTCTGATACAGACGCGTGGCAACTATCATTTCCAACTCTGGGACGA	247		
DB	173	TGTAACCGATCATGGTGTACGAAGACGATGGGCGAATTTCCATTTCCAACTCTGGGACGA	232		
QY	248	TGATTTCAATCAGTCTCTGATCTCAACGCCCTTATGGAGCACTCGATTTACCGGAAACGTGC	307		
DB	233	TGATGTCAATCAGTCT---TTACCAACCGGCTTATGAGAAATAATCGTACCTGGAGCGTGC	289		
QY	308	TGACAGACTTATTTGGGGAAGTAAAGGATATAATCTTCAATTTCAAGTCGCTGGAAGATGG	367		
DB	290	TGAGAACTGATCGGGGAAGTAAAG---AACATGTTCAATTCGATGTCATTAGAAGATGG	346		
QY	368	AG-----GCAATGATCTCTTCAACGACTTTTGTGCTGATGACGTTGA	412		
DB	347	AGAGTTAATGATCGCGTCAATGATCTCAITCAACGCTTTTGGATTTGCGACAGCCTGA	406		
QY	413	ACGTTTGGGAATCGACAGGCAATTTCAAAAAGAGATAAAACGCACTCGAATATGTTAA	472		
DB	407	ACGTTTGGGATCCATAGACATTTCAAAAGATGAGATAAAATCGCGCTTGATTTATGTTA	466		
QY	473	CAGTTATTGGAACGAAAAGGCAATGGATGTGGAGGAGAGTGTGTGACTGACCTCAA	532		
DB	467	CAGTTATTGGGCGAAATGCGATCGATCGGGAGGAGAGTGTGTACTGATCTGAA	526		
QY	533	CTCAACGCGCTTGGGGCTTCCAACTCTCCGACTACCGGATACACTGTGCTTTCAGATGT	592		
DB	527	CTCAACTCGCTTGGGGCTTCCAAACCCCTACGACTACCGGATACCCGCTGTCTTCAGATGT	586		
QY	593	TTTGAACGTTTTTAAAGCAAAAATGGGCAATTTTCTCCACTGCCAATATTTCAGATAGA	652		
DB	587	TTTCAAGCTTTTCAAAAGGCGCAAAATGGGCAAGTTTCTGCTCTGAAAATATTTCAGACAGA	646		
QY	653	GGGAGAGATTAGAGGCGTCTCAATTTATTCAGGGCTTCCCTCGTCCGCTTTCCCGGCGA	712		
DB	647	TGAAGAGATCAGAGGCGTCTCGAATTTATTTCCGGGCTTCCCTCATTCGCTTTCCAGGGGA	706		
QY	713	GAAAGTTATGGATGAAGCTGAAACATTTCTCTACAAAATATTTAAGAGAGAGCCCTGCAAAA	772		
DB	707	GAAAATTTATGATGAGSCTGAAATCTCTCTACCAAATATTTTAAAGAGAGCCCTGCAAAA	766		
QY	773	GATTCGGCAATCCAGTATCTTTCACTAGAGATACGGGAGCTTCTCGAATATGTTGGCA	832		
DB	767	GATTCGGGCTCCAGT---CTTTCCGGAGAGATTCGGGAGCTTTTGGAAATATGTTGGCA	823		

Qy	833	CACCAATTTGGCAGCGCTTGGAAAGCAAGAAATTACATGGACGCTTTTGGACACGACACTAA	892		
Db	824	CACATATTTGGCGGATTTGGAAAGCAAGAAATTACATCCAAGTCTTTTGGACAGGACTGA	883		
Qy	893	AAATGAAGACGG-----CGCCGAGAAACTTTTGAACCTTGCAAAATTTGGAGTTCAA	943		
Db	884	GAACAGAAATCATATGTGAAGAGCAAAAACCTTTTAGAACTCGCAAAATTTGGAGTTCAA	943		
Qy	944	TATATTTCACTCTTTACAAAGAGAGAGCTTAAACATGTTTCCCGATCGTGAAGAGACTC	1003		
Db	944	CATCTTTCAATCTTTTACAAAAGAGGAGTTAGAAAGTCTGGTCAGATCGTGAAGAAATC	1003		
Qy	1004	GGGTTCTCTGAGATGACCTTCTGTGACATCGTCACGTGGAATACTACGCTTTGCTTC	1063		
Db	1004	GGGTTTCTCTGAGATGACCTTCTGCGGACATCGTCACGTGGAATACTACCTTTGGCTTC	1063		
Qy	1064	CTGATTTGCGTTTGGAGCTTCAACATTTCTGGATTCAGACTCGGCTTTGCAAGAGCTGCA	1123		
Db	1064	CTGATTTGCGTTTGGAGCTTCAACATTTCTGGATTCAGACTCGGCTTTGCAAGAGCTGCA	1123		
Qy	1124	TCCTTATCAGCGTCTTCTGACGACATGTAGACGCTCTTGGACAGTAGAGAGCTTGAAC	1183		
Db	1124	TCCTTATCAGCGTCTTCTGACGATATGTAGACACCTTTCGGCACAGTAGAGAGCTTGAAC	1183		
Qy	1184	CTTCACAGCGCAAAATTAAGAGATGGGATCCGCTCGCGATGGAATGCTTCCAGAAATAT	1243		
Db	1184	CTTCACAGCGCAAAATTAAGAGATGGGATCCGCTCGCGATGGAATGCTTCCAGAAATAT	1243		
Qy	1244	GAAAGGATGTPACATGATGTTTATACACCGTAAATGAAATGCTCGAGTGGCAGAGAA	1303		
Db	1244	GAAAGGATGTPACATGATGTTTATACACCGTAAATGAAATGCTCGAGAGGCGAGAGAA	1303		
Qy	1304	GGCTCAAGCGCGACACGCTCAACTATGCAAGACAGCGCTTGGAGAGGCTGTTTGTATC	1363		
Db	1304	GGCTCAAGCGCGAGATAGCTTCAATATGCTCGGGAAGCTTGGAGAGGCTTATATTTGATC	1363		
Qy	1364	GTAATATCAGGAAGCAAAAGTGGATCGGCACTGTTTATCTGCCCCACGTTTGGAGGATCT	1423		
Db	1364	GTAATATCAGGAAGCAAAAGTGGATCGGCACTGTTTATCTGCCCCACGTTTGGAGGATCT	1423		
Qy	1424	GGAGAACGGAAAGTTAGCTCTGCTCATTCGCCCATGCGCATGCAACCCATTTTGAAGCT	1483		
Db	1424	CGAAGATGGGAAGTTAGCTCTGCTCATTCGCCCATGCGCATGCAACCCATTTTGAAGCT	1483		
Qy	1484	GGACATCCCTTTCTGATCAGCTTCAAGGAGTTGACTTCCCATCGAGCTCAATGA	1543		
Db	1484	GGACATCCCTTTCTGATCAGCTTCAAGGAGTTGACTTCCCATCGAGCTCAATGA	1543		
Qy	1544	CTTGATATGATCATCTCTCGATTAAGAGGTGATACACGGTGTCTACAAGGCGAGACAGGC	1603		
Db	1544	CTTGATATGATCATCTCTCGATTAAGAGGTGATACACGGTGTCTACAAGGCGAGACAGGC	1603		
Qy	1604	CCGTGGAGAGAGCTTCTGCTATATCATGTTATATGAAAGACAAATCTCGATTTAAACGA	1663		
Db	1604	TCGTGGAGAGAGAGCTTCTGCTATATCATGTTATATGAAAGACAAATCTCGATTTAAACGA	1663		
Qy	1664	AGAAGATGCTCTCAATCATATCAACTTCATGATCAGGACGCAATCAGAGAAATTAATTTG	1723		
Db	1664	AGAAGATGCTCTCAATCATATCAACTTCATGATCAGGACGCAATCAGAGAAATTAATTTG	1723		
Qy	1724	GGAGCTTCTAAAGCCGAGACAAACAGTGTCCCATCACTTCCAGAAACACCCATTTGACAT	1783		
Db	1724	GGAACTTCTCAAAACGAGACATCAATGTTCCCATCTCGGCGAAGAAACATGCTTTGACAT	1783		
Qy	1784	AACAGAGTTTGGCTATCAGATACCGGTTACAGATACCGGTTTCCCAACGTTGA	1843		
Db	1784	CGCCAGAGCTTTCATTTAGGGCTACAAATACCGGAGCGGCTACAGCGTTTCCCAACGTTGA	1843		
Qy	1844	AACAAAGAGTTTGGCTATCAGATACCGGTTACAGATACCGGTTTCCCAACGTTGA	1903		
Db	1844	AACAAAGAGTTTGGCTATCAGATACCGGTTACAGATACCGGTTTCCCAACGTTGA	1903		

QY 1904 ATCTACAAATTAAGTGGATGCCCTATGGTGTATATAGGCGACACAAAAATAAATAT 1963  
 DB 1904 ATCTATGCCCTATGCTATGTCGGGTAAATAATATATGGAAGGTAGCGCTGGATGTAGA 1963  
 QY 1964 GTTCTGTTAGTAAAGCTGTAATTTATGAAAAAATAAAAAAAAAAAAAA 2013  
 DB 1964 GGATAGGTTGTATATTAATTAATAAGTGTAAATTAATAAAAAAAAAAAAAA 2013

RESULT 10  
 AGU87909 2018 bp mRNA linear PLN 24-SEP-1997  
 LOCUS Abies grandis pinene synthase (AG3.18) mRNA, complete cds.  
 DEFINITION U87909  
 ACCESSION U87909  
 VERSION U87909.1 GI:2411482  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Abies grandis  
 Abies grandis  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferales; Pinaceae; Abies.  
 1 (bases 1 to 2018)  
 Bohlmann, J., Steele, C.L. and Croteau, R.  
 Monoterpene synthases from grand fir (*Abies grandis*). cDNA  
 isolation, characterization, and functional expression of myrcene  
 synthase, (-)-(4S)-limonene synthase, and (-)-(1S,5S)-pinene  
 synthase  
 J. Biol. Chem. 272 (35), 21784-21792 (1997)  
 MEDLINE 97413772  
 PUBMED 9268308  
 2 (bases 1 to 2018)  
 Bohlmann, J., Steele, C.L. and Croteau, R.  
 Direct Submission  
 Submitted (31-JAN-1997) Institute of Biological Chemistry,  
 Washington State University, Clark Hall, Pullman, WA 99164-6340,  
 USA

FEATURES  
 Location/Qualifiers  
 1..2018  
 /organism="Abies grandis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:46611"  
 1..2018  
 /gene="AG3.18"  
 6..1892  
 /gene="AG3.18"  
 /note="monoterpene synthase; terpene synthase; monoterpene  
 synthase; terpene cyclase"  
 /codon\_start=1  
 /product="pinene synthase"  
 /protein\_id="AAB71085.1"  
 /db\_xref="GI:2411483"  
 /translation="MALVSTAPLAKSKLHSLISSTHELKALSRITIPALGMSRRGKS  
 IPTSISMSSTVVVDGVRMRMGDFHSLNWDVDSLOSLPAYEKSYLEAEKLGIV  
 KNMFNSLDEGLMSPLNDLQRLMIVDSLERLGHRHFKDIBKISGALDVIYVWGEN  
 GIGCRSVVDLNSALGRLTLRLHGVPSVDVFKAFQNGQFSCSENIQIDSEIR  
 GVNLFRASLIAPPGKIMDEAEIFSTKYLKALQKIPVSSLREIGDVLVEYHTYL  
 PLRANLYQVFGDNTENKSVYKSKLLEKLEFNIFOSLOKRELSLVRWKESG  
 FEMTFCRHRYEYVTLACIAFEPOHSGFRLGFAKCHLITVLDMDYDFGVDSLE  
 LFTATWKRDPSSIDCLPEYMKGVYIAVDVTNEMAREAEAGRODTLYAREAWAY  
 ILSYMEARWATGYLPSDFEYENGKVSCHGRISALQPLITMDIPEPDHILKEVDFP  
 SKLNDLACAILRLRGTRCYKADRAEGEASSISCYMKONPGVSEBDALDHINMISD  
 VIKGLNWEILLKPDINVPISAKKHAFFIARAFHYGYKYRDGYSVANVETSLVTRTLL  
 SVPL"

ORIGIN  
 Query Match 64.9%; Score 1306.8; DB 8; Length 2018;  
 Best Local Similarity 81.0%; Pred. No. 0;  
 Matches 1595; Conservative 0; Mismatches 342; Indels 33; Gaps 5;  
 QY 68 CAGTGTGCTCAGTTCTTCTCATGAGATTAAAGGCTCTCGTAGAACAATCCCAACTCT 127  
 DB 53 CAATCGTTGATCATGTTCTACCATGAGCTTAAGGCTCTCTCTAGAACAAATCCAGCTCT 112  
 QY 128 TGAATCTGCAGCGCGGGGAAATCCGTGCGGCAATCCATAAACATGTTGTTGACAAGCGT 187

DB 113 AGGAATGAGTAGGCGAGGGAATCTATCACTCTTCCATCAGCATGAGCTCTACACCGT 172  
 QY 188 CGCATCTACTGATCTGTACAGAGCGGTGGGAACTATCATTCACACCTGTGGGAGA 247  
 DB 173 TGTAAACCGATGATGGTGTACGAGACGATGGGCGATTTCCATTCACACCTCTGGGAGA 232  
 QY 248 TGATTTTCATACAGTCTCTGATCTCAACGCTTATGGAGCACCTGATTAACCGGAAACGTGC 307  
 DB 233 TGATGTCAATCACTCT---TTACCAACGCTTATGAGGAAAAATCGTACCTGAGCGTGC 289  
 QY 308 TGACAGACTTATTTGGGGAAGTAAAGATATATGTTCAATTTCAAGTCTGCTGGAAGATGG 367  
 DB 290 TGAGAAACTGATCGGGGAAAGTAAAG---AACATGTTCAATTCGATGTCTATTAGAAGATGG 346  
 QY 368 AG-----GCAATGATCTCTTCAACGACTTTTGTGCTGGTGTGATGAGCTTGA 412  
 DB 347 AGAGTTAATGATCGCGCTCAATGATCTCANTCAAGCCCTTTGGATGTGACAGCGCTTGA 406  
 QY 413 ACGTTTGGGAATCGACAGGCATTTCAAAAAAGAGATAAAAAACGGCACTCGATTATGTTAA 472  
 DB 407 ACGTTTGGGATCCATAGACATTTCAAAGATCAGATATAAATCGGCGCTTGATTATGTTTA 466  
 QY 473 CAGTTATTGGAACGAAAAAGGATTTGGATGTGGGAGGAGAGTGTGTGACTGACCTCAA 532  
 DB 467 CAGTTATTGGGGGAAAAATGGCATCGGATCGGGGAGGAGAGTGTGTACTGATCTGAA 526  
 QY 533 CTCACCGCTTGGGGCTTCGAACTCTCGGACTTACACGGATACACTGTGCTCTTCAGATGT 592  
 DB 527 CTCACCTGCTTGGGGCTTCGAACTCTAGACTTACAGGATACCCGGTCTCTTCAGATGT 586  
 QY 593 TTTGAACTGTTTTAAAGACAAAAATGGCAATTTCTCCACCTGCCAATATTTCAGATAGA 652  
 DB 587 TTTCAAAGCTTTCAAAGGCCAAAAATGGGAGTTCCTGCTCTGAAAAATATTTCAGACAGA 646  
 QY 653 GGGAGAGATAGAGGCGTTCTCAATTTATTCAGGGCTCCCTGCTCGCTTTCCCGGCGA 712  
 DB 647 TGAAGAGATCAGAGGCGTTCTGAAATTTATTCGGGGCTCTCCCTCAITTCGCTTTCCAGGGGA 706  
 QY 713 GAAAGTTATGGATGAAGCTGAAACATTTCTCTCAAAATATTAAAGAGAGCCCTGCAAAA 772  
 DB 707 GAAATATTGGATGAGGCTGAAATCTCTCTACCAANTATTAAAGAGAGCCCTGCAAAA 766  
 QY 773 GATTCGCGCATCCAGTATATTTCTACTAGAGATAACGGAGCTTCTGGAATATGTTGGCA 832  
 DB 767 GATTCGCGTCTCCAGT---CTTTCGCGAGAGATCGGGGAGCGTTTGGAAATATGTTGGCA 823  
 QY 833 CACCAATTTGCCACGCTTGGAGCAGGAATACATGGAGCTCTTTGGACAGCACACTAA 892  
 DB 824 CACATATTTTCCGCGGATTTGGAAGCAAGGAATACATCCAAAGTCTTTGGACAGCACACTGA 883  
 QY 893 AAATAAGAACGC-----CGCCGAGAAACTTTTAGAACTTGCAGAAATTTGGAATTCAA 943  
 DB 884 GAACAGAGTCTATATGTGAGAGCAAAAACTTTTAGAACTCGCAAAATTTGAGATTCAA 943  
 QY 944 TATATTTCACTCTTTCAAGAGAGAGAGTTAAACATGTTTCCGATCGTGGGAAGACTC 1003  
 DB 944 CATCTTTCAATCTCTTCAAAAGAGGAGTTAGAAAGTCTGGTCAGATGGTGGAAAGAACT 1003  
 QY 1004 GGGTCTCTCGAGATGACCTTCTGACATCGTCAGTGAATACATACGCTTTGGCTTC 1063  
 DB 1004 GGGTTTTCTTGAATGACCTTCTGCCGACATCGTCAGTGAATACATACGCTTTGGCTTC 1063  
 QY 1064 CTGCATTGCTTCGAGCCTCAACATTTCTGATTCAGACTCGGCTTTACCAAGATGTCCTCA 1123  
 DB 1064 CTGCATTGCTTCGAGCCTCAACATTTCTGATTCAGACTCGGCTTTTCCCAAGACGTGTCA 1123  
 QY 1124 TCTTATCACGGTCTTTCAGCAGATGACGCTTTCGGGACAGTACACGAGCTTGGAACT 1183  
 DB 1124 TCTTATCACGGTCTTTCAGCAGATGATACGACACCTTCGGCAGAGTACGAGCTTGGAACT 1183  
 QY 1184 CTTTCACAGCGACAAATTAAGAGATGGGATCCGTCCGCGATGGAATGCTTCAGAAATATAT 1243

Db	1184	CTTCACAGCGACAATGAAGAGATGGGATCGTCTCGATAGATGGCTTCCAGATATAT	1243
QY	1244	GAAGAGAGTGTACATGATGTTTATCACACCGTAATAATGAATGGCTCGAGTGCAGAGAA	1303
Db	1244	GAAGAGAGTGTACATGATGTTTATCACACCGTAATAATGAATGGCTCGAGTGCAGAGAA	1303
QY	1304	GGCTCAAGCGCCGAGACACCGCTCAACTATCAACAGACAGGCTGGGAGCGCTGTTTGATTC	1363
Db	1304	GGCTCAAGCGCCGAGATACGCTCATATGCTCGGGAAGCTTGGGAGGCTTATTTGATTC	1363
QY	1364	GTATATGCGAGGAAGCAAGTGGATCGCCACTGGTATCTGCGCCACGTTTGAGGAGTACTT	1423
Db	1364	GTATATGCGAGGAAGCAAGTGGATCGCCACTGGTATCTGCGCCACGTTTGAGGAGTACTT	1423
QY	1424	GGAGAACGGGAAGTGTAGCTCTGCTCATCGCCCATGCGCACTGCGAACCCATCTGACGTT	1483
Db	1424	CGAGAAATGGGAAGTGTAGCTCTGCTCATCGCCCATGCGCACTGCGAACCCATCTGACAA	1483
QY	1484	GGACATCCCTTTCCTGATACATCTCAAGGAAGTGTAGCTCTGCGAACCCATCTGACAA	1543
Db	1484	GGACATCCCTTTCCTGATCATATCTCAAGGAAGTGTAGCTCTGCGAACCCATCTGACAA	1543
QY	1544	CTTGATATGTAATCATCTTCGATTAAAGAGGTGATACACGGTCTCAAGGCGACAGAGGC	1603
Db	1544	CTTGATATGTAATCATCTTCGATTAAAGAGGTGATACACGGTCTCAAGGCGACAGAGGC	1603
QY	1604	CCGTGGAGAAAGCTTGTCTATATCATGTTATATGAAAGCAATCTCTGGATTAAACGA	1663
Db	1604	TCGTGGAGAAAGCTTGTCTATATCATGTTATATGAAAGCAATCTCTGGATTAAACGA	1663
QY	1664	AGAAGATGCTCTGATCATATCAATCTCATGATCAGGAGCGCAATCAGAGATTAATG	1723
Db	1664	GAAGATGCTCTGATCATATCAAGCGCATCATGATCAGGAGCGCAATCAGAGATTAATG	1723
QY	1724	GGAGTCTTAAAGCGAGCAACAGTGTTCCTCATCTCCAGAAACACGCAATTTGACAT	1783
Db	1724	GGAGTCTTAAAGCGAGCAACAGTGTTCCTCATCTCCAGAAACACGCAATTTGACAT	1783
QY	1784	AAGCAGAGTGTGGCATCAGGTTACAGATACCGAGATGGCTTACAGCTTTGCCAAGTTGA	1843
Db	1784	CGCCAGAGCTTTCATACAGGCTCAAAATACCGAGCGCTTACAGCTTTGCCAAGTTGA	1843
QY	1844	AACAAAGAGTGTGGTATGAGAACCGTTCATGAACTGTGCTTGTAAACACACTTCAA	1903
Db	1844	AACAAAGAGTGTGGTATGAGAACCGTTCATGAACTGTGCTTGTAAACACACTTCAA	1903
QY	1904	ATCTACAAATATTAAGTGGATGCGCTTATGATATAGGCGACACAAAAATAAATAT	1963
Db	1904	ATCTATGCGCTATGCTATGTCGGGTAAATATATGTTGGAAGTATGCGGTTGATGTA	1963
QY	1964	GGTTGTGTTAGTAAAGCTGTAATTTATGAAAAAATAAATAAATAAATAAATAAATAA	2013
Db	1964	GGATAAGTGTGTTAATTTAATAAAGTGTGTAATTTAATAAATAAATAAATAAATAAATAA	2013

RESULT 11  
 LOCUS AY237645  
 DEFINITION Picea sitchensis pinene synthase mRNA, complete cds.  
 ACCESSION AY237645  
 VERSION AY237645.1  
 KEYWORDS GI:34582666  
 SOURCE Picea sitchensis (Sitka spruce)  
 ORGANISM Picea sitchensis  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.  
 1 (bases 1 to 1958)  
 AUTHORS Byun McKay, S.A., Hunter, W.L., Goddard, K.A., Wang, S.X., Martin, D.M., Bohlmann, J. and Plant, A.L.  
 TITLE Insect Attack and Wounding Induce Traumatic Resin Duct Development and Gene Expression of (-)-Pinene Synthase in Sitka Spruce  
 JOURNAL Plant Physiol. 133 (1), 368-378 (2003)  
 PUBMED 12970502

REFERENCE  
 AUTHORS Byun McKay, A.S., Hunter, W.L., Goddard, K.A., Wang, S.X., Martin, D., Bohlmann, J. and Plant, A.L.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-FEB-2003) Biology, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada  
 FEATURES  
 Location/Qualifiers  
 source 1..1958  
 /organism="Picea sitchensis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3332"  
 /clone="19-2-1/12-3"  
 75..1958

CDS

/note="terpene synthase; monoterpene synthase; terpene cyclase; (-)-pinene synthase"  
 /codon\_start=1  
 /product="pinene synthase"  
 /protein\_id="AAP72020.1"  
 /db\_xref="GI:34582667"  
 /translation="MALVSVAPMASRCLHKSLSAHELKTIPTILGMSRSGKS  
 ATPMSMLTITVSDGQVOREMGDPHSNLWDDFTQSLSTSYGSEYRERAEILGEV  
 KKMNSMSSEGBELISPHNDLIQRWVWDSVERLIGIERHFKNEIKSALDYVYVMSK  
 GIGRESVADLNSLTALFRLRLHGVASDVNLNFKDQGFACSPQTEEIRLS  
 VINLYRASLIAPGCEKMEERAEIFSAYLEBSLQKISVLSQETIRDLVYEWHTYLP  
 RMEARHIDVFQDQTKNSKINTEKLLEAKLEFNI PHSLKRELEVLTVMWDXSGS  
 PQMTFCRRHRYVYTLASCIAPFQHSGRFLGFAKACHILITLDDMYDTFGTVDELEL  
 PTAAMKRPDPAADCLPEYMGVLLILYDVTNETSREAEKAGQRTLDYARAMDYLL  
 DSYMQEAKWATGYLPTFAEYENGKVSSEHRTSALQPLIMDIPPHILIKREDFPS  
 KLNDASAILRLGRTICYKADRRAGEASSISCMKNDPNEGATEDALDHINAMISDV  
 IRLGNWELLNPNVSPISKKHVFDIRAPHYGYKRDGYSVANETKSLVRITVDP  
 VTL"

ORIGIN

Query Match 63.3%; Score 1273.6; DB 8; Length 1958;  
 Best Local Similarity 81.8%; Pred. No. 0;  
 Matches 1574; Conservative 0; Mismatches 299; Indels 51; Gaps 7;

QY	8	GTGCTTCTTATCTGATAGCAAGCTGAAATGGCTCTCTTCTTATCTTCTTATCTCCGCTGTTTC	67
Db	47	GTGATCAATATTTGATATTCAGCAGGATGCTGCTCAGAGCTTAAGACGATCTGAT	106
QY	68	CAGTCTGCTGCT-----CAGTCTTCTCATGAGATTAAGGCTCTCCGCTAG	112
Db	107	CAGATCATGCTGACAAATCGTTGAGCAGTCTGCTCAGAGCTTAAGACGATCTGAT	166
QY	113	AACAATCCCACTCTCGAATCTCGACGCGGGAATCCGTCGCAATCCCAATAAAT	172
Db	167	AACAATCCCACTCTGGAATGCTAGGCGAGGAAATCTCGACGCTCTCATGAGCAT	226
QY	173	GTGTTTGACAAGCGCTCGCATCTACTGATCTGTACAGAGACGCGTGGCAACTATCATTC	232
Db	227	GAGTTTGACCACTCTGATCTGATGATGTTGACAAAGCGATGGCGATTTCCATTC	286
QY	233	CAACCTGTGGACGATGATTTTATGATCTCTGATCTCAACGCTTATGGAGCACCTGA	292
Db	287	CAACCTATGGAACGATGATTTTATGATCTCTCAACGCTTATGGAGCACCTTC	343
QY	293	TTACCGGGAACGCTGACAGACTTATTTGGGAAGTAAGATATATGTTCAATTTCAA	352
Db	344	TTACCGGGAACGCTGACAGACTGATTTGGGAAGTAAG---AAGATGTTCAATTTCAAT	400
QY	353	GTGCTGGAAGATGGAG-----GCAATGATCTCTTCAACGACTTTTGTCT	397
Db	401	GTATCGGAAGACGAGAGTAAATCAGTCCCAACATGATCTCATTCACGCGTTTGGAT	460
QY	398	GGTCGATGAGCTTGAACGTTTGGGAATTCGACAGGCAATTTCAAAAAAGAGATAAAACGCG	457
Db	461	GGTTGATGTTGTTGAACGTTTAGGATAGAGACATTTCAAAAAATGAGATAAAATCAGC	520
QY	458	ACTCGAATTTGTTAAACAGTTATTGGAACGCAATTTGATGTTGGAGGAGGATGT	517
Db	521	GCTGGATTATGTTTACAGTTATTGGAGCGGAAAGGCAATTTGGATGTGGAAGAGAAACGCT	580

Qy	518	TGTGACTGACCTCAACTCAACCGCCTTTGGGCTTTGGA	CTCCGACTACACGGATACAC	577
Db				
Db	581	TGTTGCTGATCTGAACTCAACTGCTTTGGGGTTTCGAA	CTCTTCGACTACACGGATACGC	640
Qy				
Qy	578	TGTGCTCTCAGATGTTTTGAA	CGTTTTTAAAGACAAAATGGCAATTTTCTCCACTGCG	637
Db				
Db	641	TGTGCTCGCCGACGTTTTGATCTTTTCAAGACCAAAATGGCAGTTTGCATGCTCTCC	700	
Qy				
Qy	638	CAATATTTCAGATAGAGGAGAGATTAGAGCGTTCTCAA	TTTTATTTCAGGGCCTCCCTCGT	697
Db				
Db	701	---GAGTCAGACAGAGAGATCAGAGCGTACTCAATTTATATCGGGCTTCCTCAT	757	
Qy				
Qy	698	CGCCTTTCCCGCGAGAAAGTTATGGATGAAGCTGAAA	CATTCTCTA	757
Db				
Db	758	CGCCTTTCCCGGAGAAAAGTTATGGAAGAGGCTGAAATCTTCTCTG	CGCAAAATATTTGGA	817
Qy				
Qy	758	AGAAGCCCTGCAAAAGATTCCGGCATCCAGTATAC	TTTTCACATAGACATACGGAGCTTCT	817
Db				
Db	818	AGAATCCCTGCAAAAGATTTCA	GTTCCAGT---CTTTACAGGAGATACGGACGTTCT	874
Qy				
Qy	818	GGAATATGGTTGGCACCAAAATTTGCCACGCTTTGGAAGCAAGGAATTA	CATGGAAGCTTT	877
Db				
Db	875	GGAATATGGTTGGCACACATATTTGCCAAGATGGAAGCAAGGAATCA	CATCGAGCTCTT	934
Qy				
Qy	878	TGGAAGACACATAAAAAATAAGAACGC-----CGCGAGAAA	ACTTTTAGAATTGC	928
Db				
Db	935	TGGAAGACACACTCAGAAATCAAAAGTCAATGATAAAATACAGAAA	AAAACTTTTGGAACTTGC	994
Qy				
Qy	929	AAAAATTGGAATTCAAATATATTTCACTCCCTTACAGAGAGAGAGCTTAA	AAATCTTTTCCCG	988
Db				
Db	995	AAAAATTGGAGTTCAACATCTTTCACTCCCTTACAAAAGAGAGAGTTAGA	ATATCTCGTAG	1054
Qy				
Db	1055	ATGTTGGAAGATTCAGGTTCACTCAATGACCTTTGTTCGACATCGTCA	CGTCAATTA	1114
Qy				
Db	1049	CTACGCTTTGGCTTCTCTGATGAGATGACCTTCTGTTCGACATCGTCA	CGTCAATTA	1108
Db				
Db	1115	CTACACTTTTGGCTTCTGATGCTTTGAGCTTCAACATTTCTGGATTCAG	ACTCGGCTT	1174
Qy				
Qy	1109	TACCAAGATGCTCATCTTATCAAGGTTCTTGACGACATCTACGAGCT	CTTCGGCACAGT	1168
Db				
Db	1175	TGCCAAAGCGTGTATATTAATCAAGATTCGAGCATATGTA	CGACACTTTTGGAACTG	1234
Qy				
Db	1169	AGACGAGCTGGAATCTTTCAAGCGCAAAATTAAGAGATGGGATCCGT	TCGCGATGGAATG	1228
Db				
Db	1235	TGACGAGCTCGAACTCTTTCAGCGGCAATGAAGAGATGGATCCGT	TCGCGCTGATG	1294
Qy				
Qy	1229	CCTTCCAGAAATATGAAAGAGTGTTACATGATGGTTTATCA	CACCGTAAATGAATGGC	1288
Db				
Db	1295	CCTTCCAGAAATATGAAAGAGTGTTACCTGATACTTTTACGACACCGT	TAATGAAACGTC	1354
Qy				
Qy	1289	TCGAGTGCAGAGAGGCTCAAGCCGAGACACCTCAACTATGCAAGAC	CAGGCTTGGGA	1348
Db				
Db	1355	TCGAGAGCAGAGAGGCTCAAGGAAGAGACACGCTCGCATGCTCGA	CGGGCTTGGGA	1414
Qy				
Qy	1349	GGCGTGTTTGATTGCTATATGCAAGAGCAAAAGTGGATCGCCACT	CGGTTATCTGCCAC	1408
Db				
Db	1415	CGATTATCTTGATTCGATATGCAAGAGCAAGTGGATCGCCACT	GTTTATCTGCCAAC	1474
Qy				
Qy	1409	GTTTGAAGAGTACTTTGGAGAACCGGAAAGTTAGCTCTGCTCAT	CGCCCATGCGCACTGCA	1468
Db				
Db	1475	GTTTGGGAGTACTACGAGAACCGGAAAGTTAGCTCTGGTCA	TCGCACTCGGCATTCGA	1534
Qy				
Qy	1469	ACCGATTCGAGTTGGACATCCCTTTCTTGATCACATCTCAAGGAAG	TTGACTTTCCC	1528
Db				
Db	1535	ACCGATTCGAGATGGACATCCCTTTCTTCTCTCATCTCAAGGAG	TTGACTTTTCC	1594
Qy				
Qy	1529	ATCGAAGCTCAATGACTTGAATGATCATCTTTCGATTAAGAGGT	GATACACGGTGCTA	1588
Db				
Db	1595	ATCGAAGCTCAATGACTTGGCATCTGCCATCTCTCGACTACGAGGT	GATACCGGGTGCTA	1654
Qy				
Qy	1589	CAAGGCAGACAGGCGCGTGGAGAGAGAGCTTCGCTATATCATG	TATATGAAGACAA	1648

Db	1655	C	AGGCGGACAGAGCCGTCGAGAGAAGCTTCTAGTATATCTTGTGTATATGAAGACAA	1714
Qy	1649	T	CTCTGGATTAAACGAGAGAATGCTCTGATCATATCACTTTCATCATCAGGACGCAAT	1708
Db	1715	T	CTCTGGAGCACACAGAGGAAGATGCTCTCGACCATATCAACGCCATGATCAGTGATAT	1774
Qy	1709	C	CAGAGAATTAATTTGGGAGCTTCTTAAAGCCAGACAAACAGTGTTCCTCATCTTCCCAAGAA	1768
Db	1775	C	CAGAGGATTAATTTGGGAGCTTCTCAACCCAAACAGCAGTGTTCCTCATATCTTCCCAAGAA	1834
Qy	1769	A	ACAGCATTTTGCATAAAGCAGAGTTTGGCATTCAGGTTACAGATACCGAGATGCGTACAG	1828
Db	1835	A	CAATGTTTTTGCATAAAGCAGAGCTTTCCTATTCAGGTTACAAATATCGAGATGCGTACAG	1894
Qy	1829	C	TTTTCGCAACGTTTCAACCAAGAGTTTGGTGATCAGAACCGTCAATGAACTGTGCGCTTT	1888
Db	1895	C	CGTTGCCAACACANTGAAACAAAGATTTGGTGAGGAGAACCGTGATTGATCTGTGACTTT	1954
Qy	1889	G	ATAA 1892	
Db	1955	G	ATAA 1958	
RESULT 12				
AF543527				
LOCUS	AF543527	2082 bp	mRNA	linear
DEFINITION	Pinus taeda (-)-alpha-pinene synthase mRNA, complete cds.			
ACCESSION	AF543527			
VERSION	AF543527.1	GI:28894481		
KEYWORDS				
SOURCE	Pinus taeda (loblolly pine)			
ORGANISM	Pinus taeda			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 2082)			
AUTHORS	Phillips,M.A., Wildung,M.R., Williams,D.C., Hyatt,D.C. and Croteau,R.			
TITLE	cDNA isolation, functional expression, and characterization of (+)-alpha-pinene synthase and (-)-alpha-pinene synthase from loblolly pine (Pinus taeda): Stereocontrol in pinene biosynthesis			
JOURNAL	Arch. Biochem. Biophys. 411 (2), 267-276 (2003)			
MEDLINE	22510022			
PUBMED	12623076			
REFERENCE	2 (bases 1 to 2082)			
AUTHORS	Phillips,M.A., Wildung,M.R., Williams,D.C. and Croteau,R.B.			
TITLE	Direct Submission			
JOURNAL	Submitted (09-SEP-2002) Institute of Biological Chemistry, Washington State University, PO Box 646340, Pullman, WA 99164-6340			
FEATURES	USA			
source	Location/Qualifiers			
	1. .2082			
	/organism="Pinus taeda"			
	/mol_type="mRNA"			
	/db_xref="taxon:3352"			
	32. .1921			
	/note="monoterpene synthase"			
	/codon_start=1			
	/product="(-)-alpha-pinene synthase"			
	/protein_id="AA061225.1"			
	/db_xref="GI:28894482"			
	/translation="MSPSVSIVLPSDLCLPTSFIDRSREILPIHITIPNVAMRQGLI LMTASMNLTAVSDDAVIRRRDGFHSNLWDDLIOQLSPGCPGVRFAERLII EYKNSFMNSNEDGESITPLDDLIQRLWVDSVERLIDRHKFKIKSALDHVRYM KVGICGRCSEVYTDLNSGALGRLTLRHGYDVSADVLNHNKQSGCFACLTKQEDQ RVLNLVRSALIFAGPCKVMDRAESFSAKYLKALOKT PVSFSREIGVLVGYWTH LPRLEARNIDVFGDPTKNSKSYMTEKLELAKLEFNI FHAKRELEYLVYRWKKG GSPQWTCFHREVEYTLASCITAFRPOSGPRLGFAPAKAHIITVLDDMYDTFTGLDE ELFTSAIKWDSFATCECPYMKGVYMLVNTVNMESQADKAQGRDNLNYCRQWAE YIDATWQERAKWLTASGEVTPFYRNGKVSCHRVSAQLPILTTDIPPEHVLKWDI PQSLNDLSAILRLRGDTRCYQADRARGEASCI SCYMKDNFGTTTEEDALNHLNMI DVIKGLNLELLPNSVPTSAKKAHFDISRAFHCGYKKGDDGVSANIEFTKSLVKRTI DPVTL"			
CDS				



(+)-alpha-pinene synthase and (-)-alpha-pinene synthase from  
loblolly pine (*Pinus taeda*): Stereocontrol in pinene biosynthesis  
Arch. Biochem. Biophys. 411 (2), 267-276 (2003)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

2 (bases 1 to 2162)  
Phillips, M.A., Wildung, M.R., Williams, D.C. and Croteau, R.B.  
Direct Submission  
Submitted (09-SEP-2002) Institute of Biological Chemistry,  
Washington State University, PO Box 646340, Pullman, WA 99164-6340,  
USA

FEATURES

source

1..2162  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/db\_xref="taxon:3352"  
59..1942  
/note="monoterpene synthase; Pti10"  
/codon\_start=1  
/product="alpha-terpineol synthase"  
/protein\_id="AA061227.1"  
/db\_xref="GI:28894486"

CDS

/translation="MDLISVLPASASKSCVCLHKLSSSTHKLKPFCKTIRILVPRRW  
EPARMSLSSTVASEDDIORRTGGYLSNLDNDVIFLSTPYGELAYRERARLLDEV  
RDI FMSLSLEDEFSDLRLQRLWVNVERLGRHFKNEIKSALDYVSYWSEKIGIC  
GKSIITNLNLTALGERILRHGYPSADVLKHPNOLGQFVSCPSETEDIRIMVNL  
YKASLIAPFPGKVEEAESEFSEKILKLETLQKIPDCSLSGREGLVLEHGHNLPL  
PLEARNYIDVQDQTKNMPNRTKLELAKLEFNI POSIOKTELLESLRWNDSGS  
POITFRHREYVYTLASCIAPFQPSGRLGFAKACHILTVLDDMDLFGTVDELKL  
FTAIKRPDPSATDCLPOYMKGIYMMVNTVNMSEAOKAQORDTINYARQAWEDCL  
DSHMOEAKWIRFLPTFEYVLENGKVSAAHRSVALOPMLTMDIPPPHILKEVDPSP  
NUNDLACMLRLGRTRCYQADARBEETSCISCYMKDMPGATEBALNHLNVLGVP  
IKELNWEILLKPNSSVPISSKKNFIDITRAFHIGYKYRDSYSSVETKSLVWRTLLPL  
VPL"

ORIGIN

Query Match 57.4%; Score 1155.6; DB 8; Length 2162;  
Best Local Similarity 79.0%; Pred. No. 3.6e-308;  
Matches 1464; Conservative 0; Mismatches 354; Indels 36; Gaps 6;  
QY 81 AGTTCTTCTCATGAGATTAAAGGCTCTCCGTAGAACCAATCCCAACTCTTGGAAATCTGCAGG 140  
DB 125 AGTTCTTACTATAAATTAAGCTTTCTGCAAAACAATCCGGATCTTGTAAATGCCGAGG 184  
QY 141 CCGGGAAATCCGTGCGGATTCATTAACAATGTGTTTGAACGCTCGCATCTACTGAT 200  
DB 185 CGATGGAAATTCGCCAGCGCTCCATGA-----GTCGTAGTACCGTTGCATCTGAAGAT 238  
QY 201 TCTGTACAGACGCGTGGGCAACTATCATTTCAACCTGTGGACGATGATTTCAATACAG 260  
DB 239 GATATACAAAGACGACGCGCGGTTATCTTTCCAACTGTGGACGATGATGATACAG 298  
QY 261 TCTCTGATCTCAACGCCCTTATGGAGCACCTGATTACCGGAAACGTGTGACAGACTTATT 320  
DB 299 TTCTCTG--TCAACGCCCTTATGGGAACTCGCTTACCGTGAACGTGTGAGCACTGATT 355  
QY 321 GGGGAAGTAAAGATATAATGTTCAATTTCAAGTCGTGGAGATGAGGCAATGATCTC 380  
DB 356 GATGAAGTAAAGGACATATTCAAGTTCGATGTCTTGGAAAGATGAGAAATTCAGTGAATCTC 415  
QY 381 CTTTCAACGACTTTTGTGCTGCGATGAGCTTTGAACGTTTGGGAATFCGACAGGCAATTTCAAA 440  
DB 416 ATTCAACGCCCTTTGGATGGTTCGATAAGTTGAACGTTTGGGATTCGATAGGCAATTTCAAA 475  
QY 441 AAAGAGATAAAAACGCGACCTGATATTGTTAAACGTTTATGGAAACGAAAGGCAATTTGGA 500  
DB 476 AATGAGATCAAAATCAGCGCTTGATTATGTTTACAGTTTACTTGGAGCGAAAAGGCAATTTGGA 535  
QY 501 TGTGGGAGGAGATGTTGTGACCTGACCTCAACTCAACCGCTTGGGGCTTCGAACTCTC 560  
DB 536 TGTGGGACAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 595  
QY 561 CGACTACAGGATACACTGTGTCTTTCAGATGTTTGAACGCTTTTAAAGACAAAAATGGG 620

596 CGAATTACACGGATACCCCTGTTTCTGCAGACGTTTGAACAACATTTTCAGAAACCAAAATGGG 655  
QY 621 CAATTTTCTCCACTGCCAATATTTCAGATAGAGGAGAGATTAGAGGCGTTCCTCAATTTA 680  
DB 656 CAGTTTGTATCATGTCC---TAGTGACACAGAGGAGACATAGAATCATGGTCAATTTA 712  
QY 681 TTCAGGGCTCCCT-----CGTCGCTTTCCCGCGAGAAAGTTATGATGAAGAA 728  
DB 713 TATCGGGCTTCCCTCATTTGCCCTTTCCCGTTCGCTTTCCCGGTGAGAAAGTTATGGAAGAG 772  
QY 729 GCTGAAACATTTCTTACAAAATATTAAAGAGAGCCCTGCAAAAGATTCCGCGATCCAGT 788  
DB 773 GCAGAAAGCTTCTCAGAGAAATATTAAAGAGAACTCTGCAAAAGATTCCCGACTGCAGT 832  
QY 789 ATACTTTTCACTAGACATACGGGACGTTCTGGAATATGTTTGGCACACCAATTTGCCACGC 848  
DB 833 ---CTTTCAAGAGAGATAGGGGACGTTTGGAACTGTTGGCAACATGTTGGCACACAAATTTGCCGGA 889  
QY 849 TTGGAAGCAAGGAATTATCATGACGCTTTTGGACAGCACACTAAATAAAGAACGC--- 904  
DB 890 TTAGAAGCAAGGAATTATCATGACGCTTTTCGGACAAAGACACTAAGAATATGGAGGCAAAAT 949  
QY 905 -----GGCGAGAAACTTTTAGAACTTGCAGAACTGGAGTTCAACATCTTTCAATCCATA 959  
DB 950 AGGAAGACGAGAAACTGTTAGAATCTGCGAACTGGAGTTCAACATCTTTCAATCCATA 1009  
QY 960 CAAGAGAGAGCTTTAAACCATGTTTCCGATGCTGGAAGAGACTCGGGTTTCTCCTGAGATG 1019  
DB 1010 CAGAAGACAGAGTTAGAGTCCCTCTTGAGATGGTGAATGATTCAGTTTCGCTCAGATT 1069  
QY 1020 ACTCTTGTGACATCGTCACGTCGGAATACAGCTTTGGTTCCTGGAATGAGTTGGTTTCGAG 1079  
DB 1070 ACTTTCCTCGACATCGCCACGTCGAGTACTACACTTTGGCTTCTGCTGCTGCTGCTGAG 1129  
QY 1080 CTTCAACATTTGAGTTGAGCTCGCTTTTACCAAGATCTCTCATCTTATCATCGCTTCTT 1139  
DB 1130 CTTCAACATTTGAGTTGAGCTCGCTTTTGGGAAAGCTTGTCTATATCTCATGTTCTC 1189  
QY 1140 GAGGACATGTAGGAGCTTTTGGCACAGTAGACGCTGGAACTCTTTCACAGCGCAAAAT 1199  
DB 1190 GAGGATGTATGACGCTTTTCGGACAGTTTGGAGCTCAAACTTTTCACAGCGCAAAAT 1249  
QY 1200 AAGAGATGGGATCCGTCGCGATGGAATGCTTCCAGAAATATATGAAGAGTGTATCATG 1259  
DB 1250 AAGAGATGGGATCCGTCGCGCACAGATGCTTCCCAATATATGAAGAAATTTACATG 1309  
QY 1260 ATGCTTTTATCACACCGTAAATGAATGGCTCGAGTGGCAGAGAGCTCAAGCCCGAGAC 1319  
DB 1310 ATGCTTTTACAAACCGTAAATGAATGCTCTGCGGAGGCAAGAGGCTCAAGCCCGAGAC 1369  
QY 1320 ACCTCAACTATGCAAGACAGAGCTTGGGAGCGTGTGTTTGTATTCGTATATGCAAGAGCA 1379  
DB 1370 ACTCTCACTATGCTCGACAGGCTTGGAGGATGTTGCTTGTATTCACATATGCAAGAGCA 1429  
QY 1380 AAGTGGATCCGCTACTGTTATCTGCCCACTGTTGAGGAGTACTTGGAGAGCGGAAAGTT 1439  
DB 1430 AAGTGGATCCGCTACTGTTATCTGCCCACTGTTGAGGAAATCTTGGAGAGCGGAAAGTT 1489  
QY 1440 AGCTCTGCTCATCGCCCATCGCACTGCAACCCATCTGACGTTGGACATCTCCCTTTCCCT 1499  
DB 1490 AGCTCTGCTCATCGCGTATCGGATTTGCAACCCATGCTGACGATGGACATCCCTTCCCT 1549  
QY 1500 GATCAGATCTCAAGGAAGTTGACTTCCCATCGAAGCTCAATGACTTGTATGATCATC 1559  
DB 1550 CTTCACTCTCAAGGAAGTTGACTTCCCATCGAAGCTCAATGACTTGTATGATCATC 1609  
QY 1560 CTTGATTAAGAGTGTATACCGTGTCTACAGGACACAGGCGCCCTGTTGAGAGAGAGCT 1619  
DB 1610 CTTGATTAAGAGTGTATACCGTGTCTACAGGACACAGGCGCCCTGTTGAGAGAGAGCT 1669  
QY 1620 TCGTCTATATCATGTTATATGAAGACAACTCCTGGATTAACGGAAGAGATGCTCTGAAT 1679

Db 1670 TCCTGTATATCTTGTTATATGAAGACAATCTCTGGAGCACAGAGAAATGCTCTTAAT 1729

QY 1680 CATATCAACTCATGATCATGGGACGCAATCAGAGAAATTAATTTGGAGCTTCTTAAAGCCA 1739

Db 1730 CATCTCAACGTCATGATCATGGGCGTAATTAAGAATTAATTTGGAGCTTCTTCAAAACC 1789

QY 1740 GACAAACAGTGTTCCTCATCTTCCAAAGAACACGCAATTTGACATAGCAGAGTTTGGCAT 1799

Db 1790 ACAGCAGTGTGCCATTTCTTCCAAAGAAATTAATTTTGGACATTAACAGAGCTTTCCAT 1849

QY 1800 CACGGTTACAGATACCGAGATGGCTACAGCTTTGCCAAGCTTTGAAACAAAGAGTTTGGTG 1859

Db 1850 TAGGGTTACAAATACCGAGATGGCTACAGGTTTCCAGGTTTCAACAAAGAGTTTGGTG 1909

QY 1860 ATGAGAACCGTCATTGAACCTGTGCTTTGTATCAACACACTTCAAAATCTACAATA 1913

Db 1910 ATGAGAACCCCTCTTGAACCTGTGCTTTTATAACCAAAATCTAGAAATTTACAAGA 1963

RESULT 14

AP461460

LOCUS AP461460 2186 bp mRNA linear PLN 27-MAR-2003

DEFINITION Picea abies (+)-3-carene synthase (JF67) mRNA, complete cds.

ACCESSION AP461460

VERSION AP461460.1 GI:29293033

KEYWORDS

SOURCE Picea abies (Norway spruce)

ORGANISM Picea abies

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.

REFERENCE 1 (bases 1 to 2186)

Axelrod, D., Martin, D., Miller, B., Rawat, S., and Bohlmann, J. Traumatic resin defense in Norway spruce (Picea abies): methyl jasmonate-induced terpene synthase gene expression, and cDNA cloning and functional characterization of (+)-3-carene synthase Plant Mol. Biol. 51 (1), 119-133 (2003)

JOURNAL MEDLINE 22490501

PUBMED 12602896

REFERENCE 2 (bases 1 to 2186)

Faeldt, J., and Bohlmann, J. Direct Submission

AUTHORS Submitted (20-DEC-2001) Biotechnology Laboratory, University of British Columbia, 237-6174 University Blvd, Vancouver, BC V6T1Z4, Canada

JOURNAL

FEATURES

Location/Qualifiers

1..2186

/organism="Picea abies"

/mol\_type="mRNA"

/db\_xref="taxon:3329"

1..2186

/gene="JF67"

/note="PAJF67"

1..42

/gene="JF67"

43..1926

/gene="JF67"

/function="catalyzes the cyclization of geranyl phosphate to (+)-3-carene"

/codon\_start=1

/product="(+) -3-carene synthase"

/protein\_id="AA073863.1"

/db\_xref="GI:29293034"

/translation="MSVSIILPLASKSLYSKLSSTHELKALCRPIATLGMRRGKS VMASKSLITAVSDGQRRIGDHNSLNDDNFQSLSSPYGASSYGERAERLIGEV KEIENSIRTDGELVSRVLDLLOHLSDVNDVERLIGIDRHFQTEIKVSLDYVYSWSEK GIGSRDITVDLNTALGFRILRLRHGYTFVDFVDFPKDQMGRIACSDNHTEROISS ILNLFASLIAFPCKEMEAEIFSATYKLEALOTIPVSLSQEIQVLYVLRHNSLP RLEARTYDILQENTKQMLDVTNKVLELAKLEFNIPHSIQNELKSVSRWKESEGF PDLNFIHRHVEFTLVSGIDMEPKCTFRLSFVPMCHLITLVDMDVDFTGIDELRL FTAARKWDPSTBCLPEYMGVTVLYETVNEMAQKQSGRDTLSYVRQALEAYI GAYHKEAIESSGLYPTDFEYFENGKYSGRHIAITLQPTFMDLIPFPHVQLQIDFES KFNDFACSILRLRGDTRCYQADRARGEASCSICMYKDNPGSTOEDALNHNMIET I.KLNLWELLKPDNNVPISSKKHAFDINRGLHFFYNRQDGYTVASNETKNLVIKVLPEP

3' UTR

ORIGIN

Query Match 54.8%; Score 1103.2; DB 8; Length 2186;

Best Local Similarity 76.1%; Pred. No. 1.1e-293;

Mismatches 1434; Conservative 0; Mismatches 420; Indels 30; Gaps 5;

QY 64 TTTCAGGTCGCGCTCAGTCTCTCTCATGAGATTAAGGCTCTCCGTAGACAAATCCCAA 123

Db 86 TTTCAGGTCGCGCTCAGTCTCTCTCATGAGATTAAGGCTCTCCGTAGACAAATCCCAA 145

QY 124 TTTCAGGTCGCGCTCAGTCTCTCTCATGAGATTAAGGCTCTCCGTAGACAAATCCCAA 183

Db 146 CACTTGAATGTAGGCGAGGGAATCTGTCTATGGCTTCCAGAGCAGAGTTTGACCA 205

QY 184 GCGTCGATCTACTGATTTCTGTACAGAGCGGTGGGCAACTATCATTTCCAACTGTGG 243

Db 206 CTGCTGATCTGATGATGGTGTACAAAGCAGCATAGGCGATCATCATTTCCAACTGTGG 265

QY 244 ACGATGATTTACATCAGTCTCTGATCTCAACGCCCTTATGGAGCAGCTGATTTACCGGAC 303

Db 266 ACGATGATTTACATCAGTCTCTGATCTCAACGCCCTTATGGAGCAGCTGATTTACCGGAC 322

QY 304 GTGCTGACAGACTTATGGGGAAGTAAAGGATATATGTTCAATTTCAAGTCTCGTGGAG 363

Db 323 GTGCTGACAGACTTATGGGGAAGTAAAGGATATATGTTCAATTTCAAGTCTCGTGGAG 382

QY 364 ATGGA-----GGCAATGATCTCTTCAACGACTTTTGTGTCGATGACGTTG 411

Db 383 GAGAAATAGTGAGTCACGTCGATGATCTCTTCAACGACTTTTGTGTCGATGACGTTG 442

QY 412 AACGTTTGGGATCCACAGGCAATTTCAAAAAGAGATTAAGGCACTCGATTTATGTTA 471

Db 443 AACGTTTGGGATCCACAGCAATTTCAAACTGAGTAAAGTTTCCCTTGATTTATGTT 502

QY 472 ACAGTTATTTGAAAGCAAAAGGCAATTTGATGTGGGAGGAGAGTGTGTGACTGACCTCA 531

Db 503 ACAGTTATTTGAAAGCAAAAGGCAATTTGATGTGGGAGGAGAGTGTGTGACTGACCTCA 562

QY 532 ACTCAACCGCTTGGGCTTGAACCTCTCGACTACAGGATACACTGTGTCTCAGATG 591

Db 563 ACAGACTCGCTGGGTTTGAATTTCTTCGACTACATGGCTTACTCTGTGTTCCGAGC 622

QY 592 TTTTGAACGTTTAAAGACAAAATGGCAATTTTCTCCACTGCCCAATTTTCAAGTAG 651

Db 623 TGTGTGAACATTTCAAGACCAAGATGGCCCGATTTGCTT---GTTCGCAATCATACAG 679

QY 652 AGGAGAGATTAGAGCGCTTCTCAATTTATTCAGGCGCTCCCTCGTCCGCTTTCGCGCG 711

Db 680 AGAGCAGATCAGCAGCAATCTGAATTTATTTAGGCTTCTCTCATTTCTTCCGAG 739

QY 712 AGAAAGTTATGATGAAGCTCAAACTCTCTCAAAATATTTAAGAGAGCCCTGCAAA 771

Db 740 AAAAAGTTATGAAAGAGCTGAAATATTTCTCTGTACATATTTAAAGAGCCCTTACAA 799

QY 772 AGATTTCGGCATCCAGTATATTTTCACTAGAGATACGGAGCTTCTGCAATATGTTGCG 831

Db 800 CGATTCCAGTCTCCAGT---CTTTCCGCAAGAGATACAGTCTTCTCAATATGCTCGC 856

QY 832 ACACCAATTTCCCGCTTGGAGCAAGGAATTACATGGAGCTTCTTGGAGCAGCACATA 891

Db 857 ACTCAATTTGCTTAGATTTGAAGCTTAGGACTTACATCGACATACTTCAAGAGATAACA 916

QY 892 AAAATAAGA-----ACGCCCGCAGAAACTTTTAGAACTTGCATAATTTGAAATTC 942

Db 917 AAAATCAAAATGCTAGATGTGAACCAAGAAAGTTTGTAGAACTTGCATAATTTGAGTTC 976

QY 943 ATATATTTCACTCTCTTACAGAGAGAGTAAAGCAAGGAATTACATGGAGCTTCTTGGAGCAGCACATA 1002

Db 977 ATATCTCCACTCTTTTAAACAGAAATGAGTTTAAATCTGTTTCAAGATGGTGGAAAGAA 1036

[illegible]

```
QY 361 AAGATGGAG-----GCAATGATCTCTCAACAGCTTTTGTGCTGCGATG 405
Db |||||
QY 409 ACGATCGAGAAATTAATTACACCCTCAATGATCTCAATCAACGACTTTGGATGGTCGATA 468
Db |||||
QY 406 ACGTTGAACGTTTGGGAATCGACAGGCAATTTCAAAAAAGAGATAAAAAACGCGACTCGATT 465
Db |||||
QY 469 GCGTTGAACGTTTGGGGATCGATAGACATTTCAAAATGAGATAAAATCAGCGCTGGATT 528
QY 466 ATGTTAAACAGTTATTGGAACGAAAAAGGCATTGGATGTGGAGGAGAGTGTGTGACTG 525
Db |||||
QY 529 ATGTTTACAGTTATTGGAATGAAAAAGGCATTGGATGTGGAGAGATAGTGTGTGCTG 588
QY 526 ACTCAACTCAACGCGCTTTGGGCTTCGAACTCTCGACTACACGATACACTGTGTCTT 585
Db |||||
QY 589 ATCTCAATTCATCTGTTTGGGCTTCGAACTCTCGGCTGACCGGTTACACGGTGTCTCT 648
QY 586 CAGATGTTTGAACGTTTTTAAAGACAAAAATGGGCAATTTTCTCTCACTGCCAATATTC 645
Db |||||
QY 649 CAGAGGTTTTGAAGTTTTTGAAGACCAAAACGGACAGTTTGCATGCTCTCCAGTA--- 705
QY 646 AGATAGAGGAGAGATTAGAGCGGTTCTCAATTTATTCAGGCGCTCCCTCGCTGCTTTC 705
Db |||||
QY 706 AAACAGAGGGGAGATCAGAACGCTCTTAACCTTATATCGGCTTCCCTCATTCGCTTTC 765
QY 706 CCGCGCAGAAAGTTATGATGAAGCTGAAACATTTCTTACAGATACGGACGTTCTGGAATATG 825
Db |||||
QY 766 CTGGGAGAAAGTTATGACGAGCTGAAATCTTCTTCAAGATATTTGAAAGAGCGG 825
QY 766 TGCAAAAGATTCGGCATCCAGTATATCTTACTAGAGATACGGACGTTCTGGAATATG 825
Db |||||
QY 826 TGCAAAAGATTCGGACTGCACT---CTTCAAGAGATAGCCTATGCTTTGGATATG 882
QY 826 GTTGGCACACCAATTTGCCACGCTTGGAGCGAGGAATTAATGACGCTTTTGGAC--- 882
Db |||||
QY 883 GTTGGCACACCAATATGCCAAGATTTGGAAGCAAGGAATTAATGACGCTTTTGGACATC 942
Db |||||
QY 883 -----AGCACACTAAATAAGAACGCCCGCGAGAACTTTTAG 921
QY 943 CTAGTAGCCCATGGCTCAAGAGAAATAGACGCAATATATGGACGGCGAGAACTTTTAG 1002
Db |||||
QY 922 AACTTGCAAAATTTGGAATTCATATATTTCACTCTCTTACAGAGAGAGATTAAACATG 981
Db |||||
QY 1003 AACTAGCAAAATTTGGATTCAATATCTTCACTCTCTTCAAGAGAGAGATTACATATA 1062
QY 982 TTTCCGATGCTGGAAGACTCGGGTCTCTGATGACCTTCTGTGCAATCGTCACG 1041
Db |||||
QY 1063 TCTCCAGATGTTGGAAGATTCGGGTTTGGCTTAACTGGCCTTCAGTCGGCATGCTCAG 1122
QY 1042 TGAATACTACGCTTTGGCTTCTCTGCAATTTGGTTTCGAGCCTCAACATTCGGATTCAGAC 1101
Db |||||
QY 1123 TGAATACTACACTTTGGGTTCTTGCAATGCACTGACCCCAACATCGTGCAATTCAGAC 1182
QY 1102 TCGGCTTTACAGATGCTCATCTTATCACTGTTTTCGAGCATGTACGACGCTTTTCG 1161
Db |||||
QY 1183 TCGGCTTTGTCAAAACGCTGTCATCTTAACACGGTTCTTGAGCATATCTACGACATTCG 1242
QY 1162 GCACAGTAGACGCTCGAACTCTTCAAGGCAATTAAGATGGATCCGTCGCGGA 1221
Db |||||
QY 1243 GAACGATGGACGAAATCGAACTCTTCAAGAACGAGTCAGGAGATGGATCCGTCGGAGA 1302
QY 1222 TGAATGCCCTTCCAGATATATGAAAGAGTGTACATGATGGTTTATCACACCGTAAATG 1281
Db |||||
QY 1303 CAGAGAGCCTTCCAGACTATATGAAGAGTGTACATGTTACTCTACGAGCCCTAACTG 1362
QY 1282 AATGGCTCGAGTGGCAGAGAGCTCAAGCCGAGACAGCTCAACTATGCAAGACAGG 1341
Db |||||
QY 1363 AATGGCTCAAGAGCGCAGAAAAACAAGGCCGAGACAGCTCAACTATGCTGGAAGG 1422
QY 1342 CTTGGGAGGCGTGTGTTGATTCGTATATGAGGAGCAAGAGTGGATCGCACCTGGTTATC 1401
Db |||||
QY 1423 CTTGGGAGATTTATCTTTGATTCGTATATTAAGAAAGCAAGTGGATCGCACCTGGTTATC 1482
QY 1402 TGCCACGTTTGGAGACTCTTGGAGAACGGGAAAGTTAGTCTGCTCATCGCCCATGG 1461
```

```
Db |||||
QY 1483 TGCCAACTTTCAGGAGTACTTTTGAGAACGGGAAATTAGCTCTGCTTATCGCGCAGCGG 1542
QY 1462 CACTGCAACCGATTTCTGACGTTTGGACATCCCTTTCTCTGATCACATCTCTCAAGGAAGTTG 1521
Db |||||
QY 1543 CATTGACACCCATCTCTCACTTGGACGTACCGCTTCTGAAATACATCTCTGAAGGGAATTG 1602
QY 1522 ACTTCCCATCGAAGCTCAATGACTTGTATGTATCATCTCTCGATTAAGAGGTGATACAC 1581
Db |||||
QY 1603 ATTTTCCATCGAGATTCATGATTTTGGCATCTTCTCTCGACTACGAGGTGACACAC 1662
QY 1582 GTGTCTACAAGCGACAGAGCGGCCGTGGAGAAAGAGCTTCTGCTATATCATGTTATATGA 1641
Db |||||
QY 1663 GCTGCTACAAGCGGATAGGGCCGTGGAGAAAGAGCTTCTGTTATATCTTGTATATGA 1722
QY 1642 AAGACAATCTCTGGATTAAACGGAAGAGATGCTCTGAATCATATCAACTTCATGATCAGG 1701
Db |||||
QY 1723 AAGAACAATCTCTGGATCAACGGGGGAAGATGCTCTCAATCATATCAACTCCATGATCAATG 1782
QY 1702 ACGCAATCAGAAATTTAAATTTGGGAGCTTCTTAAAGCCAGACACAGCTGTTCCCATCACTT 1761
Db |||||
QY 1783 AAATATCAAGAAATTAATTTGGGAATTACTAAGACCTGTATAGCAATATTCATATGCTCTG 1842
QY 1762 CCAAGAAACACGCAATTTGACATACAGAGATTTGGCATCACCGTTTACAGATACCGAGATG 1821
Db |||||
QY 1843 CGAGGAAACATGCTTTTGACATACAGAGCTCTCCACCCTCTATAATACCGAGATG 1902
QY 1822 GCTACAGCTTTGCCAAACGTTGAAACAAAGAGTTTGGTGTATGAGAACCGCTCATTTGAACCTG 1881
Db |||||
QY 1903 GTTTCAGGCTTGGCCACTAAGGAAAGCAAAAGTCTGTCAGCAGAAATGCTCTTGAACCTG 1962
QY 1882 TGCCTTTGTAACAAACACTTCAATCTACA 1910
Db |||||
QY 1963 TGCCTTTGTAACAACTTATCTTTTACA 1991
```

Search completed: July 8, 2004, 15:05:51  
Job time : 8437.74 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2004, 09:46:39 ; Search time 909.973 Seconds  
(without alignments)  
9397.662 Million cell updates/sec

Title: US-10-025-145A-64

Perfect score: 2013

Sequence: 1 tttagctgctcttcttc.....aaaaaaaaaaaaaaaaaaaaa 2013

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_29Jan04:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002s:\*
- 7: geneseqn2003as:\*
- 8: geneseqn2003bs:\*
- 9: geneseqn2003cs:\*
- 10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2013	100.0	2013	4	AAF73411	Grand fir
2	1306.8	64.9	2018	2	AXX08644	Pinene sy
3	1306.8	64.9	2018	3	AAA38922	Grand fir
4	1306.8	64.9	2018	4	AAF73372	Grand fir
5	1092.4	54.3	1890	4	AAF73421	Grand fir
6	1072.4	53.3	2186	4	AAF73412	Grand fir
7	1071.6	53.2	2196	2	AXX08643	Myrcene s
8	1071.6	53.2	2196	3	AAA38927	Grand fir
9	1071.6	53.2	2196	4	AAF73371	Grand fir
10	1071.6	53.2	2205	2	AXX08663	Grand fir
11	1071.6	53.2	2205	4	AAF73391	Grand fir
12	929.4	46.2	2429	4	AAF73413	Grand fir
13	925.4	46.0	2089	2	AXX08645	Limonene
14	925.4	46.0	2089	3	AAA38938	Grand fir
15	925.4	46.0	2089	4	AAF73373	Grand fir
16	807.4	40.1	1513	3	AAA69551	Pinus rad
17	805.8	40.0	1634	3	AAA69644	Pinus rad
18	623.6	31.0	1173	3	AAA69643	Pinus rad
19	471	23.4	696	4	AAF73414	Grand fir
20	448.6	22.3	1885	2	AXX87534	Delta-sel
21	445.8	22.1	1865	3	AAA38933	Grand fir
22	445.8	22.1	1885	2	AXX87533	Delta-sel
23	445.8	22.1	1888	2	AXX87505	Grand fir

24	444.2	22.1	1885	2	AXX87532	Delta-sel
25	429.6	21.3	1967	2	AXX87513	Grand fir
26	429.6	21.3	1967	2	AXX08655	Grand fir
27	429.6	21.3	1967	4	AAF73383	Grand fir
28	427	21.2	2700	2	AAT97447	Pacific Y
29	427	21.2	2700	3	AAA38931	Yew taxad
30	391.6	19.5	1977	2	AXX87506	Grand fir
31	391.6	19.5	2424	2	AXX08654	Grand fir
32	391.6	19.5	2424	3	AAA38932	Grand fir
33	391.6	19.5	2424	4	AAF73382	Grand fir
34	391.6	19.5	2528	2	AXX87529	Grand fir
35	391.6	19.5	2528	2	AXX87531	Grand fir
36	391.6	19.5	2571	2	AXX87531	E-alpha-b
37	390.2	19.4	1785	2	AXX87504	Grand fir
38	390.2	19.4	1785	2	AXX87537	Gamma-hum
39	390.2	19.4	1785	2	AXX87535	Gamma-hum
40	390.2	19.4	1785	3	AAA38934	Grand fir
41	388.6	19.3	779	3	AAA69642	Pinus rad
42	387.8	19.3	2525	2	AXX87530	E-alpha-b
43	333.6	16.6	2861	3	AAA38937	Grand fir
44	285.8	14.2	462	3	AAA69611	Pinus rad
45	267.2	13.3	1416	2	AXX08656	Grand fir

#### ALIGNMENTS

RESULT 1  
AAF73411  
ID AAF73411 standard; cDNA; 2013 BP.  
XX  
AC AAF73411;  
XX  
DT 30-APR-2001 (first entry)  
XX  
DE Grand fir monoterpene synthase coding sequence fragment SEQ ID NO: 64.  
XX  
KW Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
KW terpinolene synthase; insect resistance; nutrition; ss.  
XX  
OS Abies grandis.  
XX  
PN WO200107565-A2.  
XX  
PD 01-FEB-2001.  
XX  
PF 24-JUL-2000; 2000WO-US020264.  
XX  
PR 26-JUL-1999; 99US-00360545.  
XX  
PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
XX  
PI Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
XX  
DR WPI; 2001-182782/18.  
XX  
P-PSDB; AAB69390.  
XX  
New nucleic acid encoding monoterpene synthases, for increasing terpene synthesis in plants, e.g. for increasing resistance to pests or for treatment of cancer.  
XX  
Claim 8; Page 147-149; 175pp; English.  
XX  
The present invention provides the protein and coding sequences of monoterpene synthases from the grand fir. These include (-)-camphene synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase and pinene synthase. The sequences can be used to produce transgenic plants expressing high levels of the enzymes, resulting in levels which are useful in protecting against and treating cancers, and to confer insect resistance on plants

SQ	Sequence	2013 BP; 591 A; 432 C; 464 G; 526 T; 0 U; 0 Other;	
	Query Match	100.0%; Score 2013; DB 4; Length 2013;	
	Best Local Similarity	100.0%; Pred. No. 0;	
	Matches 2013; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	TTTTGACGTCCTCTTATCTATAGCAAGCTGAAATGCTCTCTTCTTATTAATCTCCGC	60
DB	1	TTTTGACGTCCTCTTATCTATAGCAAGCTGAAATGCTCTCTTCTTATTAATCTCCGC	60
QY	61	TGTTTTCCAGGTCGTCCTCAGTCTCTCATGAGATTAAAGCTCTCCGTAGAAACAATCC	120
DB	61	TGTTTTCCAGGTCGTCCTCAGTCTCTCATGAGATTAAAGCTCTCCGTAGAAACAATCC	120
QY	121	CAACTCTTGGAAATCTGACGCGGGGAAATCCGTCCGCAATTCCTAFAAACATGTTTGA	180
DB	121	CAACTCTTGGAAATCTGACGCGGGGAAATCCGTCCGCAATTCCTAFAAACATGTTTGA	180
QY	181	CAAGCGTCGATCTACTGATTCGTACAGAGCGCTGGGCAACTATCATTCCTCAACTGT	240
DB	181	CAAGCGTCGATCTACTGATTCGTACAGAGCGCTGGGCAACTATCATTCCTCAACTGT	240
QY	241	GGACGATGATTTCAATACAGTCTCTGATCTCAACGCTTTATGGAGCACTGATTACCGGG	300
DB	241	GGACGATGATTTCAATACAGTCTCTGATCTCAACGCTTTATGGAGCACTGATTACCGGG	300
QY	301	AACGTCTGACAGACTTATTTGGGAAAGTAAAGATATATATTTCAATTTCAAGTCGTGG	360
DB	301	AACGTCTGACAGACTTATTTGGGAAAGTAAAGATATATATTTCAATTTCAAGTCGTGG	360
QY	361	AAGATGGAGCAATGATCTCTTCAACGACTTTTGTGTCGATGAGCTTGAACGTTTGG	420
DB	361	AAGATGGAGCAATGATCTCTTCAACGACTTTTGTGTCGATGAGCTTGAACGTTTGG	420
QY	421	GAATCGACAGGCAATTTCAAAAAGAGATAAAAAGCGCACTCGATTATGTTTAAAGATTAT	480
DB	421	GAATCGACAGGCAATTTCAAAAAGAGATAAAAAGCGCACTCGATTATGTTTAAAGATTAT	480
QY	481	GGAAAGAAAGGATTTGATGTTGGGAGGAGAGTGTGTGACTGACCTCAACTCAACCG	540
DB	481	GGAAAGAAAGGCAATTTGATGTTGGGAGGAGAGTGTGTGACTGACCTCAACTCAACCG	540
QY	541	CCTTGGGCTTCGAACTCTCCGACTACCGATACACTGTGTCTTCAAGTGTGTTGACG	600
DB	541	CCTTGGGCTTCGAACTCTCCGACTACCGATACACTGTGTCTTCAAGTGTGTTGACG	600
QY	601	TTTTTAAAGACAAAATGGGCAATTTTCTCCACTGCCAATATTCAGATAGAGGAGAGA	660
DB	601	TTTTTAAAGACAAAATGGGCAATTTTCTCCACTGCCAATATTCAGATAGAGGAGAGA	660
QY	661	TTAGAGCGTTCCTCAATTTATTCAGGCGCTCCCTCGCTGCTTCCCGGCGAGAAAGTTA	720
DB	661	TTAGAGCGTTCCTCAATTTATTCAGGCGCTCCCTCGCTGCTTCCCGGCGAGAAAGTTA	720
QY	721	TGGATGAAGCTGAAACATCTCTACAAAATTTTAAAGAGAGCGCTTCGAAAAGATCCGG	780
DB	721	TGGATGAAGCTGAAACATCTCTACAAAATTTTAAAGAGAGCGCTTCGAAAAGATCCGG	780
QY	781	CATCCAGTATATCTTCTACTAGAGATACGGGAGCTTCGGAATATGTTGGCAGACCAATT	840
DB	781	CATCCAGTATATCTTCTACTAGAGATACGGGAGCTTCGGAATATGTTGGCAGACCAATT	840
QY	841	TGCCACCGTTTGAAGCAGGAAATACATGGAGCTTTTGGACAGCACACTAAAATAAGA	900
DB	841	TGCCACCGTTTGAAGCAGGAAATACATGGAGCTTTTGGACAGCACACTAAAATAAGA	900
QY	901	ACGCGCGCGAGAACTTTTAGAATCTGCAAAATTTGGAATTCAAATATATTTCACTCCTTAC	960
DB	901	ACGCGCGCGAGAACTTTTAGAATCTGCAAAATTTGGAATTCAAATATATTTCACTCCTTAC	960
QY	961	AAGAGAGAGATTAAAAATGTTTCCGATCGGTGGAAGACTCGGGTCTCTCGAGATGA	1020
DB	961	AAGAGAGAGATTAAAAATGTTTCCGATCGGTGGAAGACTCGGGTCTCTCGAGATGA	1020

QY	1021	CCTTCTGTCGACATCGTCACTGGAATACTACGCTTTGGCTTCTTGCATTTGGAGC	1080
DB	1021	CCTTCTGTCGACATCGTCACTGGAATACTACGCTTTGGCTTCTTGCATTTGGAGC	1080
QY	1081	CTCAACATTTGGATTGAGATCGGCTTTTCAAGATGCTCTCATCTTATCAGGTTCTTG	1140
DB	1081	CTCAACATTTGGATTGAGATCGGCTTTTCAAGATGCTCTCATCTTATCAGGTTCTTG	1140
QY	1141	ACGATATGACAGCTCTTCGGCAGATGAGAGCTGGAACCTCTTCAAGCGACAATTA	1200
DB	1141	ACGATATGACAGCTCTTCGGCAGATGAGAGCTGGAACCTCTTCAAGCGACAATTA	1200
QY	1201	AGAGATGGGATCCGTCGCGATGGAATGCTTTCCAGATATATGAAAGGAGTGTACATGA	1260
DB	1201	AGAGATGGGATCCGTCGCGATGGAATGCTTTCCAGATATATGAAAGGAGTGTACATGA	1260
QY	1261	TGTTTTATCACCGTAAATGAAATGCTTCGAGTGCAGAGAGGCTCAAGGCCGAGACA	1320
DB	1261	TGTTTTATCACCGTAAATGAAATGCTTCGAGTGCAGAGAGGCTCAAGGCCGAGACA	1320
QY	1321	CGCTCAACTATGCAAGACAGGCTTTGGAGGCGTGTTCGATTCGTATATGCAAGAGCAA	1380
DB	1321	CGCTCAACTATGCAAGACAGGCTTTGGAGGCGTGTTCGATTCGTATATGCAAGAGCAA	1380
QY	1381	AGTGGATCCCACTGCTTATCTGCCACCGTTTGGAGGATCTTTGGAGAACGGGAAAGTTA	1440
DB	1381	AGTGGATCCCACTGCTTATCTGCCACCGTTTGGAGGATCTTTGGAGAACGGGAAAGTTA	1440
QY	1441	GCTCTGCTCATGCCCATGCGCACTCAACCCCTTCTGAGTTGGACATCCCTTTCCCTG	1500
DB	1441	GCTCTGCTCATGCCCATGCGCACTCAACCCCTTCTGAGTTGGACATCCCTTTCCCTG	1500
QY	1501	ATCACATCTCAAGGAGTTGACTTCCCATCGAAGCTCAATGACTTGTATGATCATCC	1560
DB	1501	ATCACATCTCAAGGAGTTGACTTCCCATCGAAGCTCAATGACTTGTATGATCATCC	1560
QY	1561	TTCCGATTAAGAGGTGATACACCGTGTCTCAAGGACAGAGGCGCCGCTGGAGAGAGCTT	1620
DB	1561	TTCCGATTAAGAGGTGATACACCGTGTCTCAAGGACAGAGGCGCCGCTGGAGAGAGCTT	1620
QY	1621	CGTCTATCATGTTTATGAAAGACAACTCTCGATTAAAGGAGAGATGCTCTGAATC	1680
DB	1621	CGTCTATCATGTTTATGAAAGACAACTCTCGATTAAAGGAGAGATGCTCTGAATC	1680
QY	1681	ATATCAACTTCATGATCAGGAGCGCAATCAGAGAAATTAATTTGGAGCTTCTTAAAGCCAG	1740
DB	1681	ATATCAACTTCATGATCAGGAGCGCAATCAGAGAAATTAATTTGGAGCTTCTTAAAGCCAG	1740
QY	1741	ACAACAGTGTCCCATCATCTTCCAAAGAAAACCGATTTTGACATTAAGCAGAGTTGGCATC	1800
DB	1741	ACAACAGTGTCCCATCATCTTCCAAAGAAAACCGATTTTGACATTAAGCAGAGTTGGCATC	1800
QY	1801	ACGTTTACAGATACCGGATGGCTTACAGCTTTTCCAACTTTGAAACAAAGAGTTTGGTGA	1860
DB	1801	ACGTTTACAGATACCGGATGGCTTACAGCTTTTCCAACTTTGAAACAAAGAGTTTGGTGA	1860
QY	1861	TGAAACCGCTCAITGAACTGTGCTTTGTAACAAACACTTCAATCTCAATATTAACCTG	1920
DB	1861	TGAAACCGCTCAITGAACTGTGCTTTGTAACAAACACTTCAATCTCAATATTAACCTG	1920
QY	1921	AGGATGCCCTATGGGTGTATATAGGCACACAAAATAAATATGTTGTTAGTAAAGC	1980
DB	1921	AGGATGCCCTATGGGTGTATATAGGCACACAAAATAAATATGTTGTTAGTAAAGC	1980
QY	1981	TGTAATTTTGAATAAAAAAAAAAAAAAAAAAAAAA 2013	
DB	1981	TGTAATTTTGAATAAAAAAAAAAAAAAAAAAAAAA 2013	

RESULT 2  
 AX08644  
 ID AX08644 standard; cDNA; 2018 BP.

XX AAX08644;  
 XX AC  
 XX 27-SEP-1999 (first entry)  
 XX DE  
 XX Pinene synthase gene.  
 XX  
 XX Myrcene synthase; limonene synthase; pinene synthase; flavour; aroma;  
 XX defense; plant seed; oil; meal; ss.  
 XX  
 XX Abies grandis.  
 XX  
 XX Key Location/Qualifiers  
 XX CDS 6..1892  
 XX FT \*tag= a  
 XX FT /product= "Pinene synthase"  
 XX PN WO9902030-A1.  
 XX  
 XX 21-JAN-1999.  
 XX PD  
 XX 10-JUL-1998; 98WO-US014528.  
 XX PF  
 XX 11-JUL-1997; 97US-0052249P.  
 XX PR  
 XX (UNIV ) UNIV WASHINGTON STATE RES FOUND.  
 XX PA  
 XX Bohlmann J, Steele CL, Croteau RB;  
 XX PI  
 XX WPI; 1999-120396/10.  
 XX DR P-PSDB; AAW85701.  
 XX  
 XX New isolated gymnosperm monoterpene synthase DNA - obtained from Grand  
 XX fir (Abies grandis), used to provide plants with modified production of  
 XX monoterpenes, e.g. myrcene, limonene or pinene.  
 XX  
 XX Claim 10; Page 74-77; 121pp; English.  
 XX  
 XX Nucleotide sequences encoding myrcene synthase, limonene synthase and  
 XX pinene synthase from Grand fir may be incorporated into any organism  
 XX (e.g. intact plant, animal, microbe), or derived cell culture that  
 XX produces geranyl diphosphate for the production of the aforementioned  
 XX enzymes or their products. The sequences when expressed in transfected  
 XX cells may also be used for the production or modification of flavour and  
 XX aroma properties, improvement of defense capability, and the alteration  
 XX of other ecological interactions mediated by myrcene, limonene, pinene,  
 XX or their derivatives. In particular they can be used for the production  
 XX of plant seeds for the extraction of oil or meal  
 XX  
 XX Sequence 2018 BP; 583 A; 431 C; 479 G; 525 T; 0 U; 0 Other;  
 XX  
 XX Query Match 64.9%; Score 1306.8; DB 2; Length 2018;  
 XX Best Local Similarity 81.0%; Pred. No. 0;  
 XX Matches 1595; Conservative 0; Mismatches 342; Indels 33; Gaps 5;  
 XX  
 QY 68 CAGCTGTGCTCAGTCTCTCTATGATTAAGCTCTCGTAGAACAAATCCCACTCT 127  
 Db  
 QY 53 CAATCGTGTGATCTCTCCCAATGAGCTTAAGGCTCTCTCTAGAACAAATCCAGCTCT 112  
 Db  
 QY 128 TGGAACTTCGAGCGCGGGGAAATCCGTGCGCATTCCTCCATAAATGTTTGAACGCT 187  
 Db  
 QY 113 AGGAATGAGTAGCGGAGGAAATCTATCACTCTCTTCCATCAGATGAGCTCTACACCGT 172  
 QY 188 CGCATCTACTGATCTGTACAGAGACGCGTGGGCAATATCAATCCAACTGTGGGACGA 247  
 Db  
 QY 173 TGTAAACCGATGATGGTGTACGAACGCGATGGGCGATTCATTCCAAACCTCTGGGACGA 232  
 QY 248 TGATTTTCATACAGTCTCTGATCTCAACGCTTATGGAGCAGCTGATACCGGHAACGTGC 307  
 Db  
 QY 233 TGATGTCATACAGTCT---TTACCACGCGCTTATGAGGAAAAATCGTACCTGGAGCGTGC 289  
 QY 308 TGACAGACTTATTTGGGGAAGTAAAGGATATAATGTTCAATTTCAAGTCGCTTGGGAAGATGG 367  
 Db

Db 290 TCAGAAACTGATCGGGAGTAAG---AACATGTTCAATTCGATGCTATTAGAAGATGG 346  
 QY 368 AG-----GCAATGATCTCTTCAACGACATTTTCTGGTGGTGCATGACCTTGA 412  
 Db 347 AGAGTTAATGAGTCGCTCAATGATCTCAATCAACGCTTTTGGATTTGCACAGCCTTGA 406  
 QY 413 AGCTTTGGGAATCGACAGCAATTTCAAAAAGAGATAAAACGGCACTCGATTTATCTTAA 472  
 Db 407 ACGTTTGGGATCCATAGACATTTCAAGATGAGATAAAATCGGCTTGTATTTATTTA 466  
 QY 473 CAGTTATTGGAAACGAAAAGGCAATTTGGATGTGGGAGGAGAGTGTGTGACTGACCTCAA 532  
 Db 467 CAGTTATTGGGGCGAAAATGGCATCGGATCGGGAGGAGAGTGTGTGTACTGATCTGAA 526  
 QY 533 CTCAACCGCTTTGGGCTTCGAACCTCTCGACATCAACGATACACTGTGCTCTCAGATGT 592  
 Db 527 CTCAACTGGTGGGCTTCGAACCTCTAGACTACACGATACCCGGGTGTCTTCAGATGT 586  
 QY 593 TTTGAACGTTTTTAAAGACAAAATGGCAATTTTCTCCCACTGCAATATTTCAGATAGA 652  
 Db 587 TTTCAAAGCTTTCAAAGGCCAAAATGGGCAGTTTTTCTCTGCTCTGAAATATTTCACACAGA 646  
 QY 653 GGGAGAGATTAGAGCGCTTCTCAATTTATTAGGGCTCTCCCTCGTGGCTTTCCCGGGA 712  
 Db 647 TGAAGAGATCAGAGCGCTTCTGAATTTATTCCGGGCTCTCCCTCATTTGCTTTCCAGGGA 706  
 QY 713 GAAAGTTATGGATGAAGCTGAAAACATTTCTACAAAATATTAAAGAGAGCCCTGCAAAA 772  
 Db 707 GAAATTTATGATGAGGCTGAAATCTTCTACCAAATATTAAAGAGAGCCCTGCAAAA 766  
 QY 773 GATTCGGCATCCAGTATATATTTTCACTAGATACGGGACGTTCTGGAATATGGTTGGCA 832  
 Db 767 GATTCGGCTCTCCAGT---CTTTGCGGAGATCGGGACGTTTGGATATATGGTTGGCA 823  
 QY 833 CACCAATTTGGCAGCTTTGGAAGCAAGAAATTACATGAGCGTCTTTGGACAGCACACTPAA 892  
 Db 824 CACATATTGCGCGGATTTGGAAGCAAGAAATTATATCCAAAGTCTTTTGGACAGGACACTGA 883  
 QY 893 AAATAAGAACGC-----CGCCGAGAACTTTTAGAACTTCGAAAATTTGGATTTCAA 943  
 Db 884 GAAACAGNAGTCATATGTGAAGAGCAAAAATTTTGAATCTCGAAAATTTGGATTTCAA 943  
 QY 944 TATATTTCATCTTTACAAGAGAGAGAGTTAAAACATGTTTCCCGATGGTGAAGAATC 1003  
 Db 944 CATCTTTCAATCTTTACAAAAGAGGAGTTAGAAAGTCTGCTCAGATGTTGGAAGAATC 1003  
 QY 1004 GGGTCTCTGAGATGACCTTCTGTGACATCGTCACTGGAATCTACGCTTTGGGCTTC 1063  
 Db 1004 GGGTTTTCTTGAGATGACCTTCTGCGGACATCGTCAGTGGAATCTACACTTTGGCTTC 1063  
 QY 1064 CTGCATTTGCGTTTCGAGCTCAACATTTCTGGATTCAGACTCGGCTTTACCAAGATGCTCA 1123  
 Db 1064 CTGCATTTGCGTTTCGAGCTCAACATTTCTGGATTCAGACTCGGCTTTTCGAAGAGCTGCA 1123  
 QY 1124 TCTTATCAACGGTTCTTCAACGACATGTACGACGCTCTTCGGCACAGTAGACGCTGGAAT 1183  
 Db 1124 TCTTATCAACGGTTCTTCAACGACATGTACGACACCTTCGGCACAGTAGACGCTGGAAT 1183  
 QY 1184 CTTTCACAGCGCAATTAAGAGATGGGATCCGTTCGGGATGGAATGCTTCAGAAATATAT 1243  
 Db 1184 CTTTCACAGCGCAATTAAGAGATGGGATCCGTTCGGGATGGAATGCTTCAGAAATATAT 1243  
 QY 1244 GAAAGGAGTGTACATGATGTTTATACACCGTAAATGAATGCTCGAGTGGGAGGAGAA 1303  
 Db 1244 GAAAGGAGTGTACATGAGCTTTACGACACCGTAAATGAATGCTCGAGGAGGAGGAG 1303  
 QY 1304 GGCTCAAGGCGGACACAGCTCAACTATGCAAGACAGGCTTTGGAGGCGTGTGTTGATTC 1363  
 Db 1304 GGCTCAAGGCGGAGATACGCTCAATATGCTCGGAAAGCTTTGGAGGCTTATATGATTC 1363  
 QY 1364 GTATATCAGGAAGCAAGTGGATCGCCACTGGTTATCTGCCACAGTTTGGAGGAGTACTT 1423  
 Db 1364 GTATATCAGGAAGCAAGTGGATCGCCACTGGTTATCTGCCCTCTCTTGTAGTACTA 1423

1424 GGAGAACGGAAAGTTAGTCTGCTCATCGCCCATCGGACCTGCAACCCATTCTGACGTT 1483  
1424 CGAGAAATGGGAAAGTTAGTCTGGTTCATCGCATATCCGCAATGCAACCCATTCTGACAA 1483  
1484 GGACATCCCTTTCTCTGATCATCTCTCAAGGAAGTTGACTTCCCATCGAAGCTCAATGA 1543  
1484 GGACATCCCTTTCTCTGATCATCTCTCAAGGAAGTTGACTTCCCATCGAAGCTCAATGA 1543  
1544 CTTGATATATATCATCTCTCGATTAAAGAGTGATACACGGTGCTACAAGGCGACAGGCG 1603  
1544 CTTGGCATGTGCATCTCTCGATTAAAGAGTGATACACGGTGCTACAAGGCGACAGGCG 1603  
1604 CGGTGAGAGAAAGCTTCTCTATATCATGTTATGAAGCAATCTGGATTACGGA 1663  
1604 TCGTGAGAGAAAGCTTCTCTATATCATGTTATGAAGCAATCTGGATTACGGA 1663  
1664 AGAAGATGCTCTGAATCATATCAATCAATGATCAGGACGCAATCAGAGAAATTAATTG 1723  
1664 GGAAGATGCTCTCGATCATATCAACGCAATGATCAGTACGTAATCAAGGATTAATTG 1723  
1724 GGAGCTTCTTAAGCCAGACAAAGTGTTCCTCATCTTCCATCACTTCCAGAAACAGCATTTGACAT 1783  
1724 GGAATCTCTCAACACAGACATCAATGTTCCATCTCGGCGAAGAAACATGCTTTTGACAT 1783  
1784 AACGAGATGTCATACAGGTTACAGATACAGATGCTACAGATGCTTCCCAACGTTGA 1843  
1784 CGCCAGAGCTTCTCCATTTACGGCTTACAAATACCGAGACGGTACAGGTTGCCAAGCTTGA 1843  
1844 AACAAAGATTTGGTGATGAGAACCGCTCATTTGAACCTGTGCTTGTGTAACACACTTCAA 1903  
1844 AACGAGATTTGGTGACGAGAACCTCTTGAATCTGTGCTTGTGTAACACACTTCAA 1903  
1904 ATCTACAATATTAACCTGAGATGCGCTATGGGTGTATATAGGGGACACAAAAATAATAT 1963  
1904 ATCTATGCCCTATGCTATGTCGGGTTAAATATATATGTAAGGTAGCGGTTGGATGTAGA 1963  
1964 GGTGTTGTTAGTAAGCTGTAATTTATGAAAAAATAAAAAA 2013  
1964 GGATAAGTTGTTTATAATTTAATAAGTTGTAATTTAAAAAATAAAAAA 2013

RESULT 3  
ID AAA38922 standard; DNA; 2018 BP.  
XX AC  
XX AC  
XX AC  
25-AUG-2000 (first entry)  
XX Grand fir pinene synthase DNA sequence SEQ ID NO:19.  
XX  
XX Synthase; protein co-ordinate data; active site; modification; terpenoid;  
XX 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;  
XX isoprenoid; breeding programme; fragrance; flavour; pheromone;  
XX defensive agent; pigment; antitumour; steroid hormone;  
XX signal transduction pathway; bile acid; affinity purification;  
XX photoreceptor; enzymatic synthesis; nutrient supplement;  
XX immunological reagent; ds.  
XX  
XX Abies grandis.  
XX  
XX WO200017327-A2.  
XX  
XX 30-MAR-2000.  
XX  
XX 17-SEP-1999; 99WO-US021419.  
XX  
XX 18-SEP-1998; 98US-0100993P.  
XX 22-APR-1999; 99US-0130628P.  
XX 23-AUG-1999; 99US-0150262P.  
XX  
XX (KENT) UNIV KENTUCKY RES DEPT.  
PA

(SALK) SALK INST BIOLOGICAL STUDIES.  
Chappell J, Manna KR, Noel JP, Starks CM;  
WPI; 2000-292839/25.  
P-PSDB; AAY90837.  
Novel terpene synthase enzymes, useful for producing terpene hydrocarbons, e.g. fragrances or antitumor agents, are derived from known enzymes by specific amino acid alterations.  
Disclosure; Page 363-366; 450pp; English.  
XX  
XX The present invention describes an isolated terpene synthase (I) comprising a region with at least 20% identity to region 265-535 of a 548 amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha-carbon atoms (alphaC) that have interatomic distances, between each other, within tabulated ranges, have a centre point (within a sphere of radius 2.3 Angstrom) within tabulated ranges, and have an ordered arrangement of R groups (defining aa side chains), excluding specific tabulated arrangements (tables given in the specification). (I), and related enzymes, are used to produce a wide range of terpenoids (e.g. cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances, flavours, pheromones, defensive agents, pigments, antitumour agents, components of signal transduction pathways, precursors of steroid hormones and bile acids, as photoreceptors and as co-factor side chains. Some synthases with little or no catalytic activity (and nucleic acids encoding them) are used as controls in the analysis of products formed by enzymatic synthesis; as nutrient supplements; for affinity purification of isoprenoids; or to develop immunological reagents or nucleic acids for monitoring expression of terpene synthase or inheritance of the gene in plant breeding programs. The new synthases may produce novel terpene products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent sequences used in the exemplification of the present invention  
XX  
XX Sequence 2018 BP; 583 A; 431 C; 479 G; 525 T; 0 U; 0 Other;

Query Match 64.9%; Score 1306.8; DB 3; Length 2018;  
Best Local Similarity 81.0%; Pred. No. 0;  
Matches 1595; Conservative 0; Mismatches 342; Indels 33; Gaps 5;  
QY 68 CAGGTGTCCTCAGTTCCTCTCATGAGTAAAGGCTCTCCGTAGAACAAATCCCACTCT 127  
DB 53 CAAATCGTTCAGTTCCTACCCATGAGCTTAAAGGCTCTCTTAGAACAAATCCAGCTCT 112  
QY 128 TGAATCTGAGCGCGGGGAAATCCGTCGCGCATTCCTATAAACATGTTGTTGACAAACGT 187  
DB 113 AGGAATGAGTAGCGGAGGAAATCTATCACTCTCTCATCAGCATGAGCTTACCACCGT 172  
QY 188 CGCATCTACTGATTTGTACAGAGACGCGTGGGCACTATCTTCCAACTCCAACTGGGACGA 247  
DB 173 TGTAAACCGATGATGGTGTACGAGACGCAATGGGCGATTTCCTTCCAACTCTGGGACGA 232  
QY 248 TGATTTTCATACAGTCTCTGATCTCAACGCTTATGGAGACCTGATTTACGGGAACGTGC 307  
DB 233 TGATGTCATACAGTCT---TTACCAACGCTTATGAGGAAATAATCGTACCTGGAGCGTGC 289  
QY 308 TGACAGACTTTATGGGGAAGTAAAGGATATAATGTTCAATTTCAAGTCGCTGGAGATGG 367  
DB 290 TGAGAAACTGATCGGGAAGTAAAG---AACATGTTCAATTCGATGTCATTAGAAGATGG 346  
QY 368 AG-----GCAATGATCTCTTCAACGACTTTTGGTGGTTCGATGAGCTTGA 412  
DB 347 AGAGTTAATGAGTCCGCTCAATGATCTCAATCAACGCTTTGGGATGTCCGACGCTTGA 406  
QY 413 ACGTTTCGGAATCGACAGGCAATTTCAAAAAGAGATAAAAACGGCACTCCGATTAATGTTAA 472  
DB 407 ACGTTTCGGAATCCATAGACATTTCAAGATGAGATAAATCGCGCTTGATTAATGTTA 466  
QY 473 CAGTTATTGGAAACGAAAAAGCATTTGATGTGGGAGGAGAGTGTGTGACTGACCTCAA 532  
DB 467 CAGTTATTGGGCGGAAAAATGCGATCGGATGCGGAGGAGAGTGTGTGACTGATCTGAA 526

QY 533 CTCAACGGCTTGGGCTTCGAACTCTCCGACTACAGGATACACTGTGTCTTCAGATGT 592  
DB 527 CTCAACTTGGTGGGCTTCGAACCCCTACGACTACAGGATACCCGGTGTCTTCAGATGT 586  
QY 593 TTTGAACGTTTTTAAAGACAAAATGGGCAATTTTCTCCACTGCGCAATATTTCCAGATAGA 652  
DB 587 TTTCAAAGCTTTCAAAGGCAAAATGGGCAATTTTCTGCTCTGAAAATATTTCCAGACAGA 646  
QY 653 GGGAGAGATTAGAGGGGTTCTCAATTTATTCAGGGGCTCCCTCGTCGCTTCCCGGCA 712  
DB 647 TGAAGAGATTAGAGGGGTTCTGAAATTTATTTCCGGGCTCCCTCATTTGCCCTTTCCAGGGGA 706  
QY 713 GAAAGTTATGGATTGAAGCTGAACAATCTCTCAAAAATTTTAAAGAGAGCCCTGCAAAA 772  
DB 707 GAAATATTAGATTAGGCTGAAATCTTCTACCAAAATTTTAAAGAGAGCCCTGCAAAA 766  
QY 773 GATTCGGGATCCAGTATATCTTTCACTAGAGATACGGGACGTTCTCGAATATGGTTGGCA 832  
DB 767 GATTCGGGCTCCAGT---CTTTCCGAGAGATCCGGGACGTTTGGAAATATGGTTGGCA 823  
QY 833 CACCATTTCACCGCTTGGAGCAAGGAATTAATGAGGACGCTTTTGGACAGCACACTAA 892  
DB 824 CACATATTTTGGCGCGATTGGAGCAAGGAATTAATGAGGACGCTTTTGGACAGCACACTGA 883  
QY 893 AAATAAGAACGC-----CGCCGAGAACTTTTAGAACCTTGCAAAATTTGGAATTTCAA 943  
DB 884 GNACACGAGTCAATATGTGAAGAGCAAAAATCTTTTAGAACCTTGCAAAATTTGGAATTTCAA 943  
QY 944 TATATTTCACTCTTACAGAGAGAGATTAAACATGTTTCCCGATGGTGAAGAGACTC 1003  
DB 944 CATCTTTCAATCTTTACAAAAGAGGAGTTGAAAGTCTGGTCAGATGGTGAAGAGATC 1003  
QY 1004 GGGTCTCTGAGATGACCTTCTGTCGACATCGTCACTGGGAATACTAGCTTTTGGCTTC 1063  
DB 1004 GGGTCTCTGAGATGACCTTCTGTCGACATCGTCACTGGGAATACTAGCTTTTGGCTTC 1063  
QY 1064 CTGATTGGTTCGAGCCTCAACATCTGGATTCAGACTCGGCTTACCAAGATCTCA 1123  
DB 1064 CTGATTGGTTCGAGCCTCAACATCTGGATTCAGACTCGGCTTACCAAGATCTCA 1123  
QY 1124 TCTTATACCGTCTTTCAGCATATAGCATCTTTCGGCACAGTACGAGCTGGAAT 1183  
DB 1124 TCTTATACCGTCTTTCAGCATATAGCATCTTTCGGCACAGTACGAGCTGGAAT 1183  
QY 1184 CTTTCAGCGCAATTAGAGATGGATCCGTCGCCGATGGAATCCCTTCCAGATATAT 1243  
DB 1184 CTTTCAGCGCAATTAGAGATGGATCCGTCGCCGATGGAATCCCTTCCAGATATAT 1243  
QY 1244 GAAAGGAGTGTACATGATGGTTTATCACACCGTAAATGAAATGGCTCGAGTGGCAGAGAA 1303  
DB 1244 GAAAGGAGTGTACATGATGGTTTATCACACCGTAAATGAAATGGCTCGAGTGGCAGAGAA 1303  
QY 1304 GGCTCAAGCGCGAGACAGCTCAACTATGCAAGACAGGCTTTGGAGGCGTGTATTC 1363  
DB 1304 GGCTCAAGCGCGAGACAGCTCAACTATGCTCGGGAAGCTTTGGAGGCGTGTATTC 1363  
QY 1364 GTATATCAGAGCAAGATGGATGGATCCGCTGTTATCTGCCACGTTTGGAGGATCTT 1423  
DB 1364 GTATATCAGAGCAAGATGGATGGATCCGCTGTTATCTGCCACGTTTGGAGGATCTT 1423  
QY 1424 GGAGAACGGGAAGTTAGCTCTGTCTATCGCCCATGCGCATCTGCAACCCATTTCTGACGTT 1483  
DB 1424 CGAATGGGAAGTTAGCTCTGTCTATCGCATATTCGCAATTCGCAACCCATTTCTGACAA 1483  
QY 1484 GGAATCCCTTTCTGATCATCTCTCAAGGAAGTTGACCTTCCCATCGAAGCTCAATGA 1543  
DB 1484 GGAATCCCTTTCTGATCATCTCTCAAGGAAGTTGACCTTCCCATCGAAGCTTAACGA 1543  
QY 1544 CTTTATATGATCATCTCTGATTAAGAGGATGATACAGGCTGCTTACAGGACAGAGGC 1603  
DB 1544 CTTTGGATGTCCTTCCATCTGATTAAGAGGATGATACAGGCTGCTTACAGGACAGAGGC 1603  
QY 1604 CCGTGGAGAAAGCTTCTGCTATATCATGTTATATGAAGACAACTCTGGAATTAACGA 1663

DB 1604 TCGTGAGAGAAGCTTCTCTATATCAATGTTATATGAAAGCAATCTCGAGTATCAGA 1663  
QY 1664 AGAAGATGCTCTCAATCATATCAACTTCATGATCAGGAGCGCAATCAGAGAATTAATTTG 1723  
DB 1664 GGAAGATGCTCTCGATCATATCAAGCCCATGATCAGTGACGTAATCAAGGATTAATTTG 1723  
QY 1724 GGAGCTTTTAAAGCCAGACAAACAGTGTTCCTCATCTTCCCAAGAAACACGCTTTGACAT 1783  
DB 1724 GGAATCTCTCAACACAGACATCAATGTTTCCCATCTCGGGAAGAAACATGCTTTGACAT 1783  
QY 1784 AAGCAGATTTGGCATACGCTTACAGATACCGAGATGCTACAGCTTTGCCAACGTTGA 1843  
DB 1784 CGCCAGAGCTTTCCATTTACGGCTACAAATACGAGACGCTACAGCTTTGCCAACGTTGA 1843  
QY 1844 AACAAAGAGTTGGTTCATGAGAACGCTCATTTCAACTGTGCTTGTGTAACAACACTTCAA 1903  
DB 1844 AACGAAGAGTTGGTTCACGAGAACCTCTCTGAAATCTGTGCTTTGTAGCAACAGCTCAA 1903  
QY 1904 ATCTCAATATTAATCTGAGGATGCGCTATGGGTGTATATAGGGCACACAAAATAAATAT 1963  
DB 1904 ATCTATGCCATGCTATGTGCGGTTAAATATATGTGGAAGTACCGCTTGGATGTAGA 1963  
QY 1964 GGTTCGTTAGTAAAGCTGTAATTTATGAAAAAATAAATAAATAAATAAATAAATAA 2013  
DB 1964 GGATAAGTTTGTATTAATTTAATAAAGTTGTAATTTAAAAAATAAATAAATAAATAA 2013

## RESULT 4

AAF73372

ID AAF73372 standard; cDNA; 2018 BP.

XX AAF73372;

XX 30-APR-2001 (first entry)

XX Grand fir (-)-pinene synthase coding sequence SEQ ID NO: 3.

XX Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;

KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;

KW terpinolene synthase; insect resistance; nutrition; ss.

XX *Abies grandis*.

XX WO200107565-A2.

XX 01-FEB-2001.

XX 24-JUL-2000; 2000WO-US020264.

XX 26-JUL-1999; 99US-00360545.

XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX Steele CL, Bohlmann J, Croteau RB, Phillips MA;

XX WPI; 2001-182782/18.

XX P-PSDB; AAB69371.

XX New nucleic acid encoding monoterpene synthases, for increasing terpene synthesis in plants, e.g. for increasing resistance to pests or for treatment of cancer.

XX Claim 38; Page 108-110; 175pp; English.

XX The present invention provides the protein and coding sequences of monoterpene synthases from the grand fir. These include (-)-camphene synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase and pinene synthase. The sequences can be used to produce transgenic plants expressing high levels of the enzymes, resulting in levels which are useful in protecting against and treating cancers, and to confer insect resistance on plants

XX	Sequence 2018 BP; 583 A; 431 C; 479 G; 525 T; 0 U; 0 Other;	
SQ	Query Match 64.9%; Score 1306.8; DB 4; Length 2018; Best Local Similarity 81.0%; Pred. No. 0; Matches 1595; Conservative 0; Mismatches 342; Indels 33; Gaps 5;	
QY	68 CAGGTGCTGCTCAGTCTCTTCATGAGATTAAGGCTCCGAGAACATCCCACTCT 127	
DB	53 CAAATCGTTCATGAGTCTACCCATGAGCTTAAGGCTCTCTCTPAGAACAAATCCAGCTCT 112	
QY	128 TGAATCTGAGGCGCGGAAATCCGTCGCGATTCCTAATAAATCATGTTTGAAGCGT 187	
DB	113 AGAATGAGTAGCGGAGGAAATCTATCACTCTCTCAATCAGATGAGCTCTACCACTG 172	
QY	188 CGCATCTACTGATCTGTACAGAGACGCTGGGCAATATCATCTCAATCAGATGAGCTCTACCACTG 247	
DB	173 TGTAAACGATGATGCTGTACGAGAGCGATGGCGGATTTCCATTCCAACCTCTGGGACGA 232	
QY	248 TGAATTCACAGCTCTGATCTCAACGCTTATGAGACCTGATTACCGGAAACGCTG 307	
DB	233 TGAATTCACAGCTCTCTCAACGCTTATGAGACCTGATTACCGGAAACGCTG 289	
QY	308 TGAACAGCTTATTTGGGGAAGTAAAGGATATAATGTTTCAATTTCAAGTCTGCTGGAAGTGG 367	
DB	290 TGAGAACTGATCGGGGAAGTAAAG---AACATGTTTCAATTCGATGTCATTAGAAGATGG 346	
QY	368 AG-----GCAATGATCTCCTTCAACGATTTTCTGCTGCTGATGACGTTGA 412	
DB	347 AGAGTTAATGAGTCCGCTCAATGATCTCAATTCGCTTATGAGGATGTCGACAGCTTGA 406	
QY	413 ACCTTTGGGAACTCGACAGGCTTTCAAAAAGAGATAAAAAACGCACTCGATTATGTTAA 472	
DB	407 ACCTTTGGGAACTCGACAGCTTTCAAAAAGAGATAAAAAACGCACTCGATTATGTTAA 466	
QY	473 CAGTTATTTGGAACGAAAAAGGCTTGGATGTTGGAGGGAGAGTGTGTTGACTGACCTCAA 532	
DB	467 CAGTTATTTGGGCGAAAAATGCGATCGGATCGGGAGGAGAGTGTGTTGACTGATCTGAA 526	
QY	533 CTCACCGCTTTGGGGCTTGGAGTCTCGAATCTCCGATACAGGATACACTGTGCTTCAGATGT 592	
DB	527 CTCACCTGCTTTGGGGCTTGGAGTCTCGAATCTCCGATACAGGATACACTGTGCTTCAGATGT 586	
QY	593 TTTGAACTTTTAAAGACAAAAATGGCAATTTCTCCACTGCCAATATTCAGATAGA 652	
DB	587 TTTCAAAGCTTTCAAAGCCCAAAATGGGAGTCTTCTGCTCTGAAATATTCAGACAGA 646	
QY	653 GGGAGAGATTAGAGCGCTTCTCAATTTATTCAGGGCTTCCCTGCTGCTTTCGGCGGA 712	
DB	647 TGAAGAGATCAGAGCGCTTCTGAATTTATTCGGGGCTTCCCTCATTTGCTTTCCAGGGGA 706	
QY	713 GAAAGTTATGATGAAGCTGAACATTTCTTCAAAATATTTAAGAGAGCCCTGCAGAAA 772	
DB	707 GAAATTTATGATGAGCTGAAATCTTCTCTCAAAATATTTAAGAGAGCCCTGCAGAAA 766	
QY	773 GATTCGGGATCCAGTATCTTTCACTAGAGATACGGGAGCTTCTGGAATATGTTGGCA 832	
DB	767 GATTCGGTCTCCAGT---CTTTCGGAGAGATCGGGAGCTTTTGGAAATATGTTGGCA 823	
QY	833 CACCAATTTGCCAGCTTGGAGCAAGAAATTAATGGAAGCTTTTGGAGAGCACTAA 892	
DB	824 CACATAATTTGCCGGATTTGAAGCAAGAAATTAATGGAAGCTTTTGGAGAGCACTAA 883	
QY	893 AAAATAAGACGC-----CGCCGAGAACTTTTGAAGCTTCGAAATTTGGAATTCAA 943	
DB	884 GAAACGAAGTCTATGTTGAAGAGCAAAAACTTTTGAAGCTTCGAAATTTGGAATTCAA 943	
QY	944 TATATTTCACTCTTACAGAGAGAGAGTTAAAAATGTTTCCCGATGCTGGAAGACTC 1003	
DB	944 CATCTTCAATCTTACAAAAGAGAGGAGTTAGAAAGTCTGGTCAAGTGGTGGAAAGATC 1003	
QY	1004 GGGTCTCTGAGATGACCTTCTGCGACATCGTCAGTGAATACAGCTTTGGCTTC 1063	

DB	1004 GGGTTTTCTGAGATGACCTTCTGCGGACATCGTCACTGGGAATATCTACACTTTGGCTTC 1063	
QY	1064 CTGCATTGGCTTCGAGCCTCAACATTTCTGGAATTCAGACTCGGCTTTTCAAGATGCTCA 1123	
DB	1064 CTGCATTGGCTTCGAGCCTCAACATTTCTGGAATTCAGACTCGGCTTTTCAAGATGCTCA 1123	
QY	1124 TCTTATACGGTTCCTGACGATGTAACAGCTCTTCCGCAAGTACAGTACAGCTGGAAT 1183	
DB	1124 TCTTATACGGTTCCTGACGATGTAACAGCTCTTCCGCAAGTACAGTACAGCTGGAAT 1183	
QY	1184 CTTTACAGCGAATTAAGAGATGGATCGCTCCGATGGAATGCTTCCAGAAATATAT 1243	
DB	1184 CTTTACAGCGAATTAAGAGATGGATCGCTCCGATGGAATGCTTCCAGAAATATAT 1243	
QY	1244 GAAAGAGTGTATCATGATGCTTATCACACCGTAAATGAAATGGCTCGAGTGGCAGAGAA 1303	
DB	1244 GAAAGAGTGTATCATGATGCTTATCACACCGTAAATGAAATGGCTCGAGTGGCAGAGAA 1303	
QY	1304 GGCTCAAGGCCGAGACACGCTCAACTATGCAAGACAGGCTTGGGAGCGGTGTTGATTC 1363	
DB	1304 GGCTCAAGGCCGAGATACGCTCACTATGCTCGGAGGCTTGGGAGGCTTATATGATTC 1363	
QY	1364 GTATATGCAAGAGCAAGTGGATCGCACTGCTTATCTGCCACCGTTTGAGGAGTACTT 1423	
DB	1364 GTATATGCAAGAGCAAGTGGATCGCACTGCTTATCTGCCACCGTTTGAGGAGTACTT 1423	
QY	1424 GGAAGACGGGAAAGTTAGCTCTGCTCATCGCCCATCGCACTGCAACCCCATCTGAGCTT 1483	
DB	1424 CGAAGTGGGAAAGTTAGCTCTGCTCATCGCACTGCAATCGCAATCGCAATCTGCAAT 1483	
QY	1484 GGAATCTCCCTTTCTGATCATCATCTTCAAGGAGTGTGACTTCCCATCGAGCTCAATGA 1543	
DB	1484 GGAATCTCCCTTTCTGATCATCATCTTCAAGGAGTGTGACTTCCCATCGAGCTCAATGA 1543	
QY	1544 CTTGATATGATCATCTTCTCGATTAAAGAGTGATACACGCTGTCTCAAGGACAGACAGGGC 1603	
DB	1544 CTTGATATGATCATCTTCTCGATTAAAGAGTGATACACGCTGTCTCAAGGACAGACAGGGC 1603	
QY	1604 CCGTGGAGAAAGAGCTTCTGCTATATCATGTTATGAAAGACAATCTCGATTAACGGA 1663	
DB	1604 TCGTGGAGAAAGAGCTTCTGCTATATCATGTTATGAAAGACAATCTCGATTAACGGA 1663	
QY	1664 AGAAGATGCTCTGATCATATCACTTATGATCAGGAGCGCATCAGAGATTAATAATG 1723	
DB	1664 AGAAGATGCTCTGATCATATCACTTATGATCAGGAGCGCATCAGAGATTAATAATG 1723	
QY	1724 GGAAGTCTTAAAGCCAGACAAAGTGTTCCTCATCTTCCCAAGAAAACAGCATTTGACAT 1783	
DB	1724 GGAAGTCTTAAAGCCAGACAAAGTGTTCCTCATCTTCCCAAGAAAACAGCATTTGACAT 1783	
QY	1784 AAGCAGAGTTTGGCATCACGTTTACAGATACCGAGATGGCTACAGCTTTGCCAACGTTGA 1843	
DB	1784 CGCCAGAGCTTTCATTACGCTTACAAATACCGAGAGCGGTACAGCGCTTGGCAACGTTGA 1843	
QY	1844 AACAAAGAGTTTGGTATGAGAACCGTCACTTGAACCTGCTTGTAAACAACTTCAA 1903	
DB	1844 AACAAAGAGTTTGGTATGAGAACCGTCACTTGAACCTGCTTGTAAACAACTTCAA 1903	
QY	1904 ATCTTACAATATTAATCTGAGGATGCCCTATGGGTGTATATATAGGGGACACAAAAATAATAT 1963	
DB	1904 ATCTTACCCCTATGCTATGTCGGGTAAATATATATGTTGGAAGGTAGCGGTGGATGTAGA 1963	
QY	1964 GGTGTGTTTAAAGCTGTAATTTATGAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 2013	
DB	1964 GGTGTGTTTAAAGCTGTAATTTATGAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 2013	

RESULT 5  
AAAF73421  
ID AAFF73421 standard; cdna; 1890 BP.  
XX  
AC AAFF73421;  
XX

30-APR-2001 (first entry)  
Grand fir monoterpene synthase coding sequence SEQ ID NO: 77.  
Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
terpinolene synthase; insect resistance; nutrition; ss.  
Abies grandis.  
WO200107565-A2.  
01-FEB-2001.  
24-JUL-2000; 2000WO-US020264.  
26-JUL-1999; 99US-00360545.  
(UNIW ) UNIV WASHINGTON STATE RES FOUND.  
Steele CL, Bohlmann J, Croteau RB, Phillips WA;  
WPI; 2001-182782/18.  
P-PSDB; AAB69393.  
New nucleic acid encoding monoterpene synthases, for increasing terpene  
synthesis in plants, e.g. for increasing resistance to pests or for  
treatment of cancer.  
Claim 18; Page 163-165; 175pp; English.  
The present invention provides the protein and coding sequences of  
monoterpene synthases from the grand fir. These include (-)-camphene  
synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
and pinene synthase. The sequences can be used to produce transgenic  
plants expressing high levels of the enzymes, resulting in levels which  
are useful in protecting against and treating cancers, and to confer  
insect resistance on plants  
Sequence 1890 BP; 550 A; 381 C; 441 G; 518 T; 0 U; 0 Other;  
Query Match 54.3%; Score 1092.4; DB 4; Length 1890;  
Best Local Similarity 76.8%; Pred. No. 3.6e-299;  
Matches 1439; Conservative 0; Mismatches 396; Indels 39; Gaps 7;  
46 TTCTATTACTCCGCTGGTTTCCAGTCTGCTCAGTTCCTCTCATGAGATTAAGGCTC 105  
26 TGTCTTCCAAATCGTCTCGCAAAATCGTGATCGTTTCTACTTATGAGCATAGGCTA 85  
106 TCCGTAGAACAAATCCCAACTCTTGGAAATCTGAGCCCGGGGAAATCCGTCGCGCAATCCA 165  
86 TCAGTAGAACAAATCCCAACTCTTGGATTCGTCGTCGAGGAGGAAATCTGTGACACATCCC 145  
166 TAAACATGTTGTGACAAACGTCGCATCTACTG---ATTCTGTACAGAGCCGTCGGCA 222  
146 TGAGATGAGTTTGTGACACCGCCAGTCTCTGATGATCATGTTGTACAAAGACGATAGTCG 205  
223 ACTATCATTCACACCTGTGGACGATGATTTTCATACAGTCTCTGATCTCAACGCTTATG 282  
206 AGTTTCATTCOAATCTGTGGACGACGATTTTCATACATCTC---TATCAACGCCCTTATG 262  
283 GAGCACTGATTAACCGGAACTGTGTGACAGACTTATTGGGGAAGTAAAGGATATAATGT 342  
263 GGGCACTTCATACCGTGAACTGTGTGATAGACTTATTGTGGAAGTAAAGGCTATA--T 319  
343 TCAATTTCAAGTCTGGAGATGAG-----GCAATGATCTCTTCAAC 387  
320 TCACCTCAATTTACGCGGAAGTGAAGAACTCAATCACTCCCTCAATGATCTCATTCAC 379  
388 GACTTTTGTGCTGATGACGTTTGAACGTTTGGGAATCGACAGGATTTCAAAAAGAGA 447  
380 GCCTTTTAATGTCGATACGTTGAACGTTTGAAGGATTTAGGATGATGACATTTCAAAAATGAGA 439

448 TAAAAACGGCACTCGATTATGTAAACAGTTATTGGAACGAAAAAGGCAATTGATGTGGCA 507  
440 TAAAGCAGCACTAGACTATGTTTACAGTTATTGGAACGAAAAAGGCAATTGCGATGGAA 499  
508 GGGAGAGTGTGTGACTGACCTCAACTCAACCGCTTGGGGCTTGGGCTTCCGACTTCCGACTAC 567  
500 GTGATAGTGGTGTGCTGATCTCAACTCACTGCCCTGGGGTTTGGAAATTTCTCGACTAC 559  
568 ACGGATACACTGTGTCTTCAGATGTTTGAACGTTTAAAGACAAAAAATGGCA--AT 624  
560 ACGGATACAGTGTCTTCAGATGTTTGAACACTTCAAGAGGAGAGAGAGGGGC 619  
625 TTTCCTCCACTGCCAATATTTCAGATAGAGGGAGAGATTAGAGCGCTTCTCAATTTATTCA 684  
620 AGTTTGTATGTTCCGCCATCCAAACAGAGGAGAGATAAAAGCGTCTCGAATTTATTTC 679  
685 GGGCTCTCTGTGCTGCTTCCCGCGAGAAAGTTATGAGTGAAGCTGAAAACATTTCTCTA 744  
680 GGGCTCTCTCTGCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 739  
745 CAAAATATTAAAGAGAGCCCTGCAAAAGATTCCCGCATCCAGTATATCTTTCACATAGAGA 804  
740 AATATATTAAAGAGAGCCCTTCAAAATATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796  
805 TACGGGACGTTGTGAATATGTTGGCACACCAATTTGCCAGCTTGGAAAGCAAGAAAT 864  
797 TAGAGTACGTTCTGGAGGATGGTTGGCAACCAATATGCCAAGATTGGAACCAAGAACT 856  
865 ACATGAGCTCTTTGGACAGCACACTAAATTAAG-----AACGCCGCCGAGAAAC 915  
857 ACATCGATGTTATGGGAGAGAACGATCGTGTAGAGACGTTATATATGAACATGGAGAAAC 916  
916 TTTTGAACCTTGCAAAATTTGAAATTTCAATATATTTTCACTTCTTCAAGAGAGAGAGTTAA 975  
917 TTTTGAATTTGCAAAATTTGAGTTTCAATATTTTCACTTCTTCAACAGAGAGAGCTAA 976  
976 AACATGTTTCCGATGTTGGAAAGACTCGGTTCTCTGAGATGACCTTCTGTCGACATC 1035  
977 AAGACCTCTCCAGATGGTGGAAAGATTGGGTTTCTCTCACCTGACATTTTCTCGGCATC 1036  
1036 GTCAGCTGGAACTACGCTTTGGCTTCTGCTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1095  
1037 GTCATGTGAAATTTACGCTCTGGCATCTTGCAATTTGAAACGATCGCAACATTCGGAT 1096  
1096 TCAGACTCGGCTTTACCAAGATGCTCATCTTATCACCGTTCTTTCAGACATGTTACGACG 1155  
1097 TCAGACTCGGCTTTGCCAAATGTGTATCTTATCACCGTTTGGACGATATATAGACA 1156  
1156 TCTTCGGCACAGTAGACGAGTGGAACTCTTTCACAGGACAAATTAAGAGATGGGATCCGT 1215  
1157 CTTTGGAAACAATGGAGAGTGGAACTCTTCACTGACGATTTAAGAGATGGGATCCGT 1216  
1216 CCGGATGGAATGCTTCCAGATATATGAAGGATGATGATGATGATGATGATGATGATGATGAT 1275  
1217 CTGCCACAGATTGCTTCCAGATATATGAAGGTTGTATCATGTTGTTTACGAAACCG 1276  
1276 TAAATGAAATGGCTCGAGTGGCAGAGAGGCTCAAGCGCGAGACACGCTCAACTATGAA 1335  
1277 TAAATGAAATTTGCTCGAGAGCAGACAGTCTCAAGCGCGAGAGACGCTCAACGATGCTC 1336  
1336 GACAGCTTGGAGAGCGGTTTGTGATTCGTATATGACAGAGCAAGAAAGTGGATTCGCCACTG 1395  
1337 GACGAGCTTGGAGAGCGCTTATCTTGATTCGTATATGAAGAGAGTGGATGGATCTCCAGTG 1396  
1396 GTTATCTCCCACTTTGAGGAGTACTTGGAGAGCGGAAAGTTAGCTCTGCTATCGCC 1455  
1397 GTTATCTGCCAAGCTTGGAGAGTACATGGAGACAGCAAAAGTTAGTTTGGTTATCGCA 1456  
1456 CATGGCACTGCAACCCATTTCTGAGTTGGAGCATCCCCCTTCTCTGATCACATCTCTCAAGG 1515  
1457 TATTGCAATTGCACCCATCTCTCACTATGATGTTCCCTTACTCACCACATCTCTGAGG 1516





Qy	413	ACGTTTGGGAATCGACAGGCATTTCAAAAAGAGATAAAAACGGCATCTCGAATTATGTTAA	472
Db	470	ACGTTTGGGGATAGCTAGACATTTCAAGAAACGAGATAACTACAGCTCTGGAATTATGTTTT	529
Qy	473	CAGTTATTGGBAACAAAAGGCATTTGGATGTGGGAGGAGAGTGTCTGACTGACCTCAA	532
Db	530	CCGTTTACTGGGAGAAACGGCATTTGGATGTGGGAGAGACAGTATTTGTTACTGATCTCAA	589
Qy	533	CTCAACCGCTTTGGGGGCTTGAACTCTCCGACTACACGGATACACTGTGTCTTCAAGTGT	592
Db	590	CTCAACTGCGTTGGGGTTTGGAACTCTTCGATTACACGGGTACACTGTATCTCCAGAGGT	649
Qy	593	TTTGAAAGTTTTTAAAGACAAAAATGGGAAATTTTCTCCACTGCCAAATATTCAGATAGA	652
Db	650	TTTAAAAAGCTTTTCAAGATCAAAATGACAGTTTTGTATGCTCCOCCC---GGTCAGACAGA	706
Qy	653	GGGAGAGATTAGAGCGGTTCTCAATTTTATTTCAGGGCTCCCTCGTCGCCCTTTCCGGCGGA	712
Db	707	GGGTGAGATCAGAAGCGTTCTTAACTTATATCGGGCTTCCCTCATTCGCTTCCCTGGTGA	766
Qy	713	GAATAGTTATGGATGAAGCTGAAACATTTCTCTCAAAAAATATTTAAGAGAAAGCCCTGAAAA	772
Db	767	GAATAGTTATGGAAAGAAAGCTGAAATCTCTTCCACAAGATATTTGAAAGAAGCTCTACAAAA	826
Qy	773	GATTCGGCATCCAGTATACTTTTCACTAGAGATACGGACGTTCTCGAAATATGTTGGCA	832
Db	827	GATTCAGTCTCCGCT---CTTTCAAGAAGATAAAGTTGTTATGGAATATGGCTGGCA	883
Qy	833	CACCAATTTGCCAGCTTGGAGCAAGAAATTAATGAGAGCTCTTTGGACAGCACACTA-	891
Db	884	CACAAATTTGCCAAGATTGGAAGCAAGAAATTTACATAGACACACTTTGAGAAAAGACACAG	943
Qy	892	-----AAATAAGAACGCCGCGAGAACTTTTAGAAGCTTCGCAAAATTTGGAATTCAA	943
Db	944	TGCATGGCTCAATAAAAATGCTGGGAAGAAGCTTTTGAAGCTTCGCAAAATTTGGAGTTCAA	1003
Qy	944	TATATTTCACTCCCTTACAAGAGAGAGAGTTTAAAAACATGTTTCCGAGCTGGTGGAAAGACT	1003
Db	1004	TATATTTAACTCCTTACACAAAAGGAATTACAATATCTTTGAGATGGTGGAAAGAGTC	1063
Qy	1004	GGGTTCTCGAGATGACCTTCTGTGCAATCGTCAGCTGGGAATPACTACGCTTTGGCTTC	1063
Db	1064	GGATTTGGCTAAATTTGACATTTGCTGGGCATCGTCATGTGGAATTTCTACACTTTGGCCTC	1123
Qy	1064	CTGCATTCGCTTCGAGCCTCAACATTTCTGGATTCAGACTCGGCTTTTACCAGATGTCTCA	1123
Db	1124	TTGTATTGCCAATGACCCAAACATTTCTGCAATTCAGACTAGGCTTCGCCAAATGTCTCA	1183
Qy	1124	TCTTATACCGGTTCTTTGACGACATGTACGAGCTTTTCGGCACAGTACAGAGCTGGAACCT	1183
Db	1184	TCTTTGTACAGTTTTTGACGAGATATTTACGACACTTTTGGAACGATTCACGAGCTTGAACCT	1243
Qy	1184	CTTTCACAGCGCAATTAAGACATGGGATTCGCTCCGCGATGGAATGCCTTCCAGAAATATAT	1243
Db	1244	CTTTCACATCTGCANTTAAGATGGAATTCATCAGAGATAGACACCTTCCAGAAATATAT	1303
Qy	1244	GAAGAGGTGTACATGATGGTTTTATCAACCGTAAATGAAATGCTCGAGTGGCAGAGAA	1303
Db	1304	GAATGTGTATACATGGTTCGTTTTGAAACTGTAAATGAACTGACAGAGAGCGCGAGAA	1363
Qy	1304	GGCTCAAGCGCCGAGACACGCTCAACTATGCAAGACAGGCTTGGGAGCGGTGTTTTGATTC	1363
Db	1364	GACTCAAGGGAGAAACATCTCAACTATGTTTCGAAGAGGCTTGGGAGGCTTATTTTGAATTC	1423
Qy	1364	GTATATGAGGAAGCAAAAGTGGATGGCCACTGGTTTATCTGCCACGTTTGGAGGATCTTT	1423
Db	1424	ATATATGAAGAGCAAAATGGATCTCTAATGGTTATCTGCCAATGTTTGAAGAGTACCA	1483
Qy	1424	GGAGAACGGGAAGTTAGCTCTGTCTCATCGCCCATGCGCACTGCAACCCCATCTTGAGGTT	1483
Db	1484	TGAGAAATGGGAAGTAGGAGCTCTGCATATTCGGGTAGCAACATTCGCAACCCCATCTCACTTT	1543
Qy	1484	GGACATCCCTTTTCTCTGATACATCCTCAAGGAAGTTGACTTCCCATCGAAGCTCAATGA	1543

Db	1544	GAATGCGATGGCTTCCTCATATACATCTTTGAAGGAATTGATTTCCATCCAGGTTCAATGA	1603
Qy	1544	CTTGATATGATCATCTTCGATTTAAGAGGTGATACACGGTGCCTACAAAGCAGACAGGGC	1603
Db	1604	TTTGGCATCGTCTTCCTTCGGCTACGAGGTGACACACGCTGCTACAAGGCCGATAGGA	1663
Qy	1604	CCGTGGAGAGAGAGCTTCGCTTATATCATGTTTATATGAAGACATCTCGATTAACGA	1663
Db	1664	TCGTGGTGAAGAAGCTTCGTGTATATCATGTTTATATGAAGACATCTCGATCAACGA	1723
Qy	1664	AGAAAGTCTCTGAATCATATCAACTTCATGATCAGGACGCAATCAGAAATTAATTTG	1723
Db	1724	AGAAGATGCCCTCAATCATATCAATGCGATGGTCAATGATATCAAGAATTAATTTG	1783
Qy	1724	GGAGCTTCTAAAGCCAGACACAGTGTCCCATCATCTTCCAAGAAACACGGATTTGACAT	1783
Db	1784	GGAATCTTCTAAGATCCAAACGACAATATCCAATGCTGCCCAAGAAACATGCTTTTGACAT	1843
Qy	1784	AAGCAGAGTTTGGCATCACCGGTTACAGATACCGAGATGGCTACAGCTTTGCCAACGTTGA	1843
Db	1844	AACAAGAGCTCTCCACCATCTCTACATATATCGAGATGGCTTAGTGTGGCCACACAGGA	1903
Qy	1844	AACAAAGAGTTTGGTGATGAGAACCCGTCATTTGAACCTGTGCCTTTGTAACAACATTTCAA	1903
Db	1904	AACAAAAAATTGGTTATGGAAACACATCTTGAATCTATGCTTTTTTAAGTAAACCATTA	1963
Qy	1904	ATCTACAAATTAATCTGAGGATGCCCTATGGGTCTATATAGGGCACA	1950
Db	1964	TCCATAATAATAAGCTCATATGCTAAATTAATTGGCCTTATGACATA	2010
RESULT 8			
AAA38927			
ID	AAA38927 standard; DNA; 2196 BP.		
AC	AAA38927;		
XX			
DT	25-AUG-2000 (first entry)		
XX			
DE	Grand fir myrcene synthase DNA sequence SEQ ID NO:29.		
XX			
KW	Synthase; protein co-ordinate data; active site; modification; terpenoid;		
KW	3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;		
KW	isoprenoid; breeding programme; fragrance; flavour; pheromone;		
KW	defensive agent; pigment; antitumour; steroid hormone;		
KW	signal transduction pathway; bile acid; affinity purification;		
KW	photoreceptor; enzymatic synthesis; nutrient supplement;		
XX	immunological reagent; ds.		
XX			
OS	Abies grandis.		
XX			
PN	WO200017327-A2.		
XX			
PD	30-MAR-2000.		
XX			
PF	17-SEP-1999; 99MO-US021419.		
XX			
PR	18-SEP-1998; 98US-0100993P.		
PR	22-APR-1999; 99US-0130628P.		
PR	23-AUG-1999; 99US-0150262P.		
XX			
PA	(KENT ) UNIV KENTUCKY RES DEPT.		
PA	(SALK ) SALK INST BIOLOGICAL STUDIES.		
XX			
PI	Chappell J, Manna KR, Noel JP, Starks CM;		
XX			
DR	WPI; 2000-292839/25.		
DR	P-PSDB; AAY90842.		
XX			
PT	Novel terpene synthase enzymes, useful for producing terpene		
PT	hydrocarbons, e.g. fragrances or anticancer agents, are derived from known		
PT	enzymes by specific amino acid alterations.		

xx PT Novel terpene synthase enzymes, useful for producing terpene  
PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known  
PT enzymes by specific amino acid alterations.

XX PS Disclosure; Page 387-390; 450pp; English.

XX The present invention describes an isolated terpene synthase (I) comprising a region with at least 20% identity to region 265-535 of a 548 amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha-carbon atoms (alphaC) that have interatomic distances, between each other, within tabulated ranges, have a centre point (within a sphere of radius 2.3 Angstrom) within tabulated ranges, and have an ordered arrangement of R groups (defining aa side chains), excluding specific tabulated arrangements (tables given in the specification). (I), and related enzymes, are used to produce a wide range of terpenoids (e.g. cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances, flavours, pheromones, defensive agents, pigments, antitumour agents, components of signal transduction pathways, precursors of steroid hormones and bile acids, as photoreceptors and as co-factor side chains. Some synthases with little or no catalytic activity (and nucleic acids encoding them) are used as controls in the analysis of products formed by enzymatic synthesis; as nutrient supplements; for affinity purification of isoprenoids; or to develop immunological reagents or nucleic acids for monitoring expression of terpene synthase or inheritance of the gene in plant breeding programs. The new synthases may produce novel terpene products. AAY38910 to AAY38938 and AAY90831 to AAY90859 represent sequences used in the exemplification of the present invention

XX SQ Sequence 2196 BP; 684 A; 416 C; 468 G; 628 T; 0 U; 0 Other;

Query Match 53.2%; Score 1071.6; DB 3; Length 2196;  
Best Local Similarity 75.0%; Pred. No. 3.1e-293;  
Matches 1430; Conservative 0; Mismatches 441; Indels 36; Gaps 6;

QY 68 CAGGTCGTGCTCAGTCTCTTCATGAGATTAAGGCTCTCCGTAGAAACAATCCCAACTCT 127  
DB 116 CAAGTCGTGTGATCAGTCAATTCATGAACATTAAGGCTCCCTATAGAACAAATCCCAACTCT 175  
QY 128 TGGATCTGCGGCGGGAATCGTCGCGAATTCATTAACATGTGTTGACAGGCT 187  
DB 176 TGGATGCGTAGGCGAGGGAATCTGTCAAGCTTCCATGAGCATGAGTTGGCCACCGC 235  
QY 188 CGCATCTACTGATCTGTACAGAGCGGTGGGCAACTATCATTCCAACCTGTGGAGGA 247  
DB 236 TGCACCTGATGATGTTGATACAAAGCGCATAGTGACTACCATTCATATCTGGAGGA 295  
QY 248 TGATTTATAGTCTCTGATCTCAACGCTTATGAGACACTGATTAACGGGAACGTCG 307  
DB 296 TGATTTATAGTCTC---TATCAACGCTTATGGGAAACCTCTTACCAAGGAACGTCG 352  
QY 308 TGACAGCTTATTTGGGGAAGTAAGGATATATGTTCAATTTCAAGTCGCTGGAAGTGG 367  
DB 353 TGAGAGATTAATTTGGGAGGTAAGAGATA---TTCAAATCAATGTACCTGGATGATGG 409  
QY 368 AGG-----CAATGATCTCTTCAACGACTTTTGTGGTTCGATGACGTTGA 412  
DB 410 AAGATTAATGAGTTCCTTTAATGATCTCATGCAACGCTTTGGATAGTCGATAGGTTGA 469  
QY 413 ACGTTTGGGAATGACAGGCAATTCAAAGAGATATAAAGCAACGCACTCGATTAATGTTAA 472  
DB 470 ACGTTTGGGGTAGCTAGACATTTCAAGAAACAGAGATAACATCAGCTCTGGATTATGTTTT 529  
QY 473 CAGTTTATTTGGAACGAAAGGATTTGGATGTGGGAGGAGAGTGTGTGACTGACCTCAA 532  
DB 530 CGTTTACTGGGAGGAAACGCGATTTGGATGTGGGAGAGACAGATTTGTTACTGATCTCAA 589  
QY 533 CTCACCGGCTTTGGGGCTTCGAACCTCTCCGACTACACGAGATACACTGTGCTCTCAGATGT 592  
DB 590 CTCNACTGGTGTGGGTTTCGAACCTCTTCGATTAACCGGTTACACTGTATCTCCAGAGGT 649  
QY 593 TTTGAACGTTTAAAGACAAAATTTGGGCAATTTTCTCCACTGCGCAATATTCAGATAGA 652  
DB 650 TTTAAAGGCTTTTCAAGATCAAAATGGAAGTTCGATGCTGCTCCCTCCCTCCCTCC 706  
QY 653 GGGAGAGATTAGAGGCTTCTCAATTTATAGGCGCTTCCCTCGTCGCTTTCCTCCGCGGA 712

DB 707 GGGTGAGATCAGAAAGCGTTCTTTAATTATATCGGGCTTCCCTCATTTGCTTCCCTGGTGA 766  
QY 713 GAAAGTTATGATGAAGCTGAAACATTTCTCTACAAAATATTTAAGAGAGCGCTTCAAAA 772  
DB 767 GAAAGTTATGAGAGAGCTGAAATCTTCTCCACAAGATATTTGAAGAGAGCTCTACAAA 826  
QY 773 GATTCCGGCATCCAGTATACCTTTTCACTAGAGATAGGAGCGTTCTTGGAAATATGTTGGCA 832  
DB 827 GATTCCAGTCTCCGCT---CTTTCAACAAGAGATAAAGTTTGTATGGAATATGGCTGGCA 883  
QY 833 CACCAATTTGCGACGCTTTGGAAGCAAGAAATACATGACGCTCTTTGGACACACACTA- 891  
DB 884 CACAAATTTGCGAGATTTGGAAGCAAGAAATACATGACACACTTTGAGAAGACACAG 943  
QY 892 -----AAAATGAAGACGCGCGAGAAACTTTTAGAACTTGCAAAATTTGGAATTCAA 943  
DB 944 TGCATGGCTCAATAAATGCTGGGAAGAGCTTTAGACTTGCAAAATTTGGAGTTCAA 1003  
QY 944 TATATTTCACTCTTACAAAGAGAGAGTTAAACATGTTTCCCGATGGTGGAAAGATC 1003  
DB 1004 TATATTTAACTCTTACAAACAAAGGAATTAACAATATCTTTTGTAGATGCTGGAAGAGTC 1063  
QY 1004 GGGTCTCTCGATGACCTTCTGTGACATCGTCACGTGGAATACTAGCTTTGGCTTC 1063  
DB 1064 GGAATTCCTAAATTTGACATTTGCTCGGCATCGTCATGTGGAATTTCTACACTTTGGCCTC 1123  
QY 1064 CTGCATTCGTTTCGAGCCTCAACATTTCTGGATTCAGACTCGGCTTTTACCAAGATGCTCA 1123  
DB 1124 TTGATTTGCCATTGACCCAAACATTTCTGCATTGAGCTTGGCCAAATTTGTGTC 1183  
QY 1124 TCTTATACCGGTTCTTACGACATGTAGACGTCTTCGGGACAGTAGAGAGCTGGAACT 1183  
DB 1184 TCTTGTCAAGTTTGGACGATATTTACGACACTTTTGGAAACGATTTGACGAGCTTGAAT 1243  
QY 1184 CTTACACGGCAATTAAGAGATGGATCGCTCGCGATGGAATGCTTCCAGAAATAT 1243  
DB 1244 CTTACATCTCGAATTAAGAGATGGAATTCATCAGAGATAGAACACCTTCCAGAAATAT 1303  
QY 1244 GAAAGAGTGTACATGATGGTTTATCACACCGTAAATGAAATGGCTCGAGTGGCAGAA 1303  
DB 1304 GAAATGTGTACATGGTTCGTTTGAACCTGTAATGAACCTGACACAGAGAGCGGAGAA 1363  
QY 1304 GGCTCAAGCGCGAGACAGCTCAACTATGGAAGACAGGCTTGGGAGGCTGTTTGTATTC 1363  
DB 1364 GACTCAAGGGAGAAACACTCTCACTATGTTTGAAGGCTTGGGAGGCTTATTTTGAATC 1423  
QY 1364 GTATATGAGGAAGCAAGTGGATCGCCACTGTTTATCTGCCACGCTTTGAGGAGTACTT 1423  
DB 1424 ATATATGGAAGAGCAAAATGGAATCTTAATGGTTATCTGCCAATGTTTGAAGAGTACA 1483  
QY 1424 GGAGAACGGGAAAGTTAGCTCTGCTCATCGCCCATGCGCACTGCAACCCATTTCTGACGTT 1483  
DB 1484 TGAGATGGGAAAGTAGCTCTGCATATCGCGTAGCAACATTTGCAACCCATCTCCTCATT 1543  
QY 1484 GGACATCCCTTTCTCTGATCAGCTCCTCAAGGAAGTTGACTTCCCATCGAGCTCAATGA 1543  
DB 1544 GAATGATGGCTTCTGATTAATCTCTGAAGGAAATTTGATTTCCATCCAGGTTCAATGA 1603  
QY 1544 CTTGATATGTATCATCTTCGATTAAGAGGTGATACACGCTGTCTACAAGGAGAGAGGCG 1603  
DB 1604 TTTGGCATCTGCTTCTTCCGCTACGAGGTGACACACGCTGTCTACAAGGCGGATAGGA 1663  
QY 1604 CCGTGGAGAAAGCTTCGCTCTATATCATGTTTATATGAAGACAAATCTCGGATTAACGGA 1663  
DB 1664 TCGTGGTGAAGAGCTTCGTATATCATGTTTATATGAAGACAAATCTCGGATCAACCGA 1723  
QY 1664 AGAGATGCTCTGATATCATATCAACTTCATGATCAGGAGCGCAATCAGAGATTAATG 1723  
DB 1724 AGAAGATCCCTCAATCATATCAATGCGCATGCTCAATGACATAATCAAGAAATTAATG 1783  
QY 1724 GGAGCTTCTAAAGCGAGACAAAGGTTTCCCATCACTTCCCAAGAAACAGCAATTTGACAT 1783  
DB 1784 GGAATCTTAAAGTCCAACGACAAATTTTCAATGCTGCGCAAGAAACATGCTTTTGACAT 1843

QY 1784 AAGCAGAGTTTGGCATCAGGTTTACAGATACCGAGATGGCTACAGCTTTGCCAAGCTTGA 1843  
 DB 1844 AACAGAGCTCTCCACCATCTCTACATATATCAGATGGCTTTAGTGTTCGCCAAGGA 1903  
 QY 1844 AACAAAGAGTTTGGTGTATGATGAGAACCGCTCATTTGAACCTGTGCTTGTAAACACATTTCAA 1903  
 DB 1904 AACAAAAAATGGTTATGGAACACACTCTTGAATCTATGCTTTTAACTATAACCATTA 1963  
 QY 1904 ATCTACAAATTAACGTAGAGATGCCCTATATGGTGTATATAGGCGACA 1950  
 DB 1964 TCCATAATAAAGCTCATATGCTAAATATTGGCCCTATGACATA 2010

RESULT 9  
 AAF73371  
 ID AAF73371 standard; cdna; 2196 BP.  
 AC AAF73371;  
 DT 30-APR-2001 (first entry)  
 XX Grand fir myrcene synthase coding sequence SEQ ID NO: 1.  
 XX Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
 KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
 KW terpinolene synthase; insect resistance; nutrition; ss.  
 XX Abies grandis.  
 OS  
 XX WO200107565-A2.  
 PN 01-FEB-2001.  
 PD  
 XX 24-JUL-2000; 2000WO-US020264.  
 PF  
 XX 26-JUL-1999; 99US-00360545.  
 PR (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 PA Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
 PI WPI; 2001-182782/18.  
 DR P-PSDB; AAB69370.  
 XX  
 PT New nucleic acid encoding monoterpene synthases, for increasing terpene  
 PT synthesis in plants, e.g. for increasing resistance to pests or for  
 PT treatment of cancer.  
 XX  
 PS Claim 28; Page 103-106; 175pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of  
 CC monoterpene synthases from the grand fir. These include (-)-camphene  
 CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 CC limonene, (-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 CC and pinene synthase. The sequences can be used to produce transgenic  
 CC plants expressing high levels of the enzymes, resulting in levels which  
 CC are useful in protecting against and treating cancers, and to confer  
 CC insect resistance on plants  
 XX  
 SQ Sequence 2196 BP; 684 A; 416 C; 468 G; 628 T; 0 U; 0 Other;  
 Query Match 53.2%; Score 1071.6; DB 4; Length 2196;  
 Best Local Similarity 75.0%; Pred. No. 3.1e-293;  
 Matches 1430; Conservative 0; Mismatches 441; Indels 36; Gaps 6;  
 QY 68 CAGGTGCTGCTCAGTTCTCTCATGAGATTAAGGCTCTCCGTAGAACAAATCCCACTCT 127  
 DB 116 CAAAGTCGTTGATCAGTTCAATTCATGAACATAAGCCCTCCCTATAGAACCAATCCAAATCT 175  
 QY 128 TCGAAATCTGAGCGCGGGGAAATCCGTCGCGCATTCATAAACATGTTTGTGACAGCGT 187  
 DB 176 TGGAAATGCGTAGCGGAGGGAATCTGTCAGCCCTTCATGAGCATCAGTTTGGCCACCGC 235

QY 188 CGCATCTACTGATTCTGTACAGAGACGCGTGGCAACTATCATTCACCACTGTGGGACGA 247  
 DB 236 TGCACCTGATGATGGTGTACAAAGACGATAGTGACTACCATTTCCAATATCTGGGACGA 295  
 QY 248 TGATTTTCATACAGCTCTGTGATCTCAACGCGCTTATGAGACACCTGATTTACCGGAAACGTGC 307  
 DB 296 TGATTTTCATACAGCTCTC---TATCAACGCGCTTATGGGAACCCCTTTACCAAGAACGTGC 352  
 QY 308 TGACAGACTTATTTGGGGAAGTAAAGGATATAATGTTCAATTTCAAGTCGCTGGGAAGATGG 367  
 DB 353 TGAGAGATTAAATTTGTGGAGGTAAAGAGATA---TTCAATTTCAATGTACCTGGATGATGG 409  
 QY 368 AGG-----CAATGATCTCCTTCAACGACTTTTTCGCTGGTGCATGACGTTGA 412  
 DB 410 AAGATTAAATGAGTTCCTTTAATGATCTCATGCAACGCTTTGGATAGTCATGAGCGTTGA 469  
 QY 413 ACCTTTGGGAATCGACAGGCAATTTCAAAAAGAGATAAAAACGCGCACTCGATTATGTAA 472  
 DB 470 ACGTTTGGGATAGCTAGACATTTCAAGAACGAGATAACATCAGCTCTGGATTATGTTT 529  
 QY 473 CAGTTATTTGAACGAAAAAGGCAATTTGGATGTGGAGGAGAGTGTGTGACTGACCTCAA 532  
 DB 530 CCGTTTACTGGGAGAAAAACGCAATTTGGATGTGGAGAGACAGTATTTGTTACTGATCTCAA 589  
 QY 533 CTCAACCGCTTGGGCTTCCGAACCTCTCGACTACACGGATACACTGTGTCTTCAGATGT 592  
 DB 590 CTCACATGCTTGGGTTTCGAACTCTTCAATTTACACGGGTACACTGTATCTCCAGAGT 649  
 QY 593 TTTTAAACGTTTTTAAAGACAAAAAATGGCAATTTTCTCCACTGCCAATAATTCAGATAGA 652  
 DB 650 TTTTAAAGCTTTTCAAGATCAAAATGGACAGTTTGTATGCTCCCCC---GGTCAGACAGA 706  
 QY 653 GGGAGAGATTAGAGCGGTTCTCAATTTATTCAGGGCTCCCTCGTGCCTTTTCCCGCGCA 712  
 DB 707 GGGTGAATCAGAAAGCGTTCTTAACTTATATCCGGGCTTCCCTCATTTGCTTCCCTGGTGA 766  
 QY 713 GAAAGTTATGGATGAAGCTGAAACATTTCTCAAAAATATTTAAGAGAACCCCTGCAGAAA 772  
 DB 767 GAAAGTTATGGAAGAGCTGAATCTTCTCCACAGATATTGAAGAGAGCTCTACAAA 826  
 QY 773 GATTCGGCATCCAGTATATCTTTCTAGAGATACGGGAGCTTCTGGAATATGTTGGCA 832  
 DB 827 GATTCAGTCTCGCT---CTTTCAAGAGATAAAGTTTGTATGGAATATGGCTGGCA 883  
 QY 833 CACCAATTTGCCACGCTTGGAGCAAGGAATTCATGAGAGCTTTTGGACAGCACACTA- 891  
 DB 884 CAAAAATTTGCCAAGATTGGAGCAAGAAATTCATAGACACACTTGGAGAAAGACACCAG 943  
 QY 892 -----AAAATAAGAACGCGCGGAGAAACTTTTAGAACTTGCAAAATTTGGAATTCAA 943  
 DB 944 TGCATGCTCAATATAAAATGCTGGGAAGAGCTTTTAGAACTTGCAAAATTTGAGTTCAA 1003  
 QY 944 TATATTTACTCTCTTCAAGAGAGAGAGTTAAACATGTTTCCGATGTTGGGAAAGACTC 1003  
 DB 1004 TATATTTAACTCTTCAACAAAAAGGAATTTACAATATCTTTTGAAGTGGTGGAAAGATC 1063  
 QY 1004 GGGTCTCTGTGATGACCTCTGTGACATCGTCACTGGAATATCTAGCTTTGGCTTC 1063  
 DB 1064 GGAATTTGCCTTAAATGACATTTGCTCGGATCGTATGGAATTTCTACACTTTGGCCTC 1123  
 QY 1064 CTGCATTTGGGTTTCGAGCTCAACATTTCTGATTCAGACTCGGCTTTTACCAAGATGCTCA 1123  
 DB 1124 TTGATTTGCCATTGATGCCAAAAACATTTCTGCAATTCAGACTAGGCTTCCGCAAAATGTGTC 1183  
 QY 1124 TCTTATCAGGTTCTTTGACGACATGTACGACGCTCTTCGGCAGATGACAGAGTGAAGT 1183  
 DB 1184 TCTTGTACAGTTTGGACGATATTTACGACATTTTGAACGATTTGACAGAGTGAAGT 1243  
 QY 1184 CTTACAGCGCAATTAAGAGATGGATCCGTCGGGATGGAATTCATCAGAGATAGAACACTTCCAGATATAT 1243  
 DB 1244 CTTACATCTGCATTTAAGAGATGGAATTCATCAGAGATAGAACACTTCCAGATATAT 1303

QY	1244	GAAGAGTGTACATGATGGTTTATACACCGTAAATGAATGGCTCGATGCGCAGAA	1303
Db	1304	GAATGTGTGTACATGTCGTGTTTGAACCTGTAATGAACCTGACACGAGCGCGGAA	1363
QY	1304	GGCTCAAGGCGGAGACACGCTCAACTATGCAAGACAGCGCTTGGGAGCGGTGTTGATTC	1363
Db	1364	GACTCAAGGAGAAACACTCTCAACTATGTTTGAAGGCTTGGAGGCTTATTTTGAATC	1423
QY	1364	GTATATCGAGGAAGCAAAAGTGGATCGGCACCTGGTTATCTGCCACGTTTGAGGAGTACTT	1423
Db	1424	ATATATGGAAGACAAATGGAATCTCTAATGGTTATCTGCCAATGTTTGAAGATACCA	1483
QY	1424	GGAGAACGGGAAAGTTAGCTCTGCTCATTCGCCCATGCGCATGCGAACCCATTTGAGCTT	1483
Db	1484	TGAGATGGGAAAGTGAAGTCTGTCATATCGCTAGCAACATGCAACCCATCCTCACTTT	1543
QY	1484	GGACATCCCTTTCCTCATCACATCTCAAGGAGTTGACTTCCCATCGAAGCTCAATGA	1543
Db	1544	GAATGCATGGCTTCCCTGATTACATCTTGAAGGGAATTTGATTTTCCATCCAGGTTCAATGA	1603
QY	1544	CTTGATATGTATCATCTCCCTCGATTGAAGAGTGTATACACGGTGTCTACAAGGCAGACAGGGC	1603
Db	1604	TTTGGCATGCTTCCCTTCCTTCGGCTACGAGGTGACACACGCTGTCTACAAGGCCGATAGGA	1663
QY	1604	CGGTGGAGAGAAAGCTTCGCTCATATCATGTTATGAAAGACAAATCCTCGATTAAACGGA	1663
Db	1664	TCGTGGTGAAGAAGCTTCGTGTATATCATGTTATGAAAGACAAATCCTGGATCAACCGA	1723
QY	1664	AGAGATGCTCTGNAATCATATCAACTTCATGATCAGGAGCGCAATCAGAGAAATTAATTTG	1723
Db	1724	AGAAGATGCCCCCTCAATCATATCAATGCCATGGTCAATGACATATCAAGAANTTAATTTG	1783
QY	1724	GGAGCTCTTAAAGCCAGACAAACAGTGTTCCTCATCACITTCCAAGAAACACGCATTTGACAT	1783
Db	1784	GGAATCTTAGATCCAAAGACATATTCAAATGCTGGCCAGAAACATGCTTTTGACAT	1843
QY	1784	AAGCAGAGTTTGGCATCACCGTTTACAGATACCGAGATGGCTTACAGCTTTGCCAACGTTGA	1843
Db	1844	AACAAGAGCTCTCCACCATCTCTACATATATCGAGATGGCTTTAGTGTTCCTCAACAGGA	1903
QY	1844	AACAAAGAGTTTGGTATGAGAACCGTCAATGAACCTGTGCTTTGTAAACACACTTCAA	1903
Db	1904	AACAAAAAATTTGGTTATGGAACACCTTCCTTGAATCTATGCTTTTTTAACTATAACCAT	1963
QY	1904	ATCTACAATATTAATCGAGGATGCCCTATGGGTGTATATAGGCACA	1950
Db	1964	TCCATAATAATAGCTCAATATGTAATTTATTTGGCCTTATGACATA	2010
RESULT 10			
AAAX08663			
ID	AAAX08663 standard; cDNA; 2205 BP.		
XX			
AC	AAAX08663;		
XX			
XX	27-SEP-1999 (first entry)		
XX	Grand Fir monoterpene synthase clone AG3.48.		
XX			
KW	Mycrene synthase; limonene synthase; pinene synthase; flavour;		
KW	monoterpene synthase; aroma; defense; plant seed; oil; meal; primer; PCR;		
KW	probe; ss.		
XX			
OS	Abies grandis.		
XX			
Key	Location/Qualifiers		
FD	57..1940		
FT	/*tag= a		
FT	/product= "Monoterpene synthase"		
XX			
XX	W09902030-A1.		
XX	21-JAN-1999.		
PD			

XX	10-JUL-1998;	98WO-US014528.	
XX	11-JUL-1997;	97US-0052249P.	
XX	(UNIW ) UNIV WASHINGTON STATE RES FOUND.		
XX	Bohlmann J, Steele CL, Croteau RB;		
XX	WPI; 1999-120396/10.		
XX	P-PSDB; AAW85710.		
XX	New isolated gymnosperm monoterpene synthase DNA - obtained from Grand fir ( <i>Abies grandis</i> ), used to provide plants with modified production of monoterpenes, e.g. myrcene, limonene or pinene.		
XX	Example 3; Page 104-107; 121pp; English.		
XX	Nucleotide sequences encoding myrcene synthase, limonene synthase and pinene synthase from grand fir may be incorporated into any organism (e.g. intact plant, animal, microbe), or derived cell culture that produces geranyl diphosphate for the production of the aforementioned enzymes or their products. The sequences when expressed in transfected cells may also be used for the production or modification of flavour and aroma properties, improvement of defense capability, and the alteration of other ecological interactions mediated by myrcene, limonene, pinene, or their derivatives. In particular they can be used for the production of plant seeds for the extraction of oil or meal		
XX	Sequence 2205 BP; 703 A; 412 C; 463 G; 627 T; 0 U; 0 Other;		
XX	Query Match	53.2%; Score 1071.6; DB 2; Length 2205;	
XX	Best Local Similarity	75.0%; Pred. No. 3.1e-293;	
XX	Matches 1430; Conservative	0; Mismatches 441; Indels 36; Gaps 6;	
Qy	68	CAGTCGTGCGCTCAGTTCCTTCATGAGATTAAGGCTCTCGTAGAACAAATCCCAATCT 127	
Db	104	CAAGTCGTGTGATCAGTTCAATTCATGAACATAAGCCTCCCTATAGAACAAATCCCAATCT 163	
Qy	128	TGGAATCTGCAGCGCGGGAAATCCGTCGGCATTCCTATAAACAATGTTTGACAAGCGT 187	
Db	164	TGGAATCGGTAGCGGAGGAAATCTGTACGCCCTCCATGAGCATCAGTTGGCCACCGC 223	
Qy	188	CGCATCTACTGATTCGTGTACAGAGACGCGTGGCACTATCATTTCCAACTGTGGGACGA 247	
Db	224	TGCACCTGATGATGGTGTACAAAGACCATAGGTGACTACCATTCCAATATCTGGGACGA 283	
Qy	248	TGATTCATACAGTCTCTGATCTCAACGCCTTATGGAGCACCTGATTTACCGGAACGTGC 307	
Db	284	TGATTTTCATACAGTCTC---TATCAACGCATTTATGGGAAACCCCTCTTACCAGGAACGTGC 340	
Qy	308	TGACAGACTTATTGGGGAAGTAAAGGATATAATGTTCAATTTCAAGTCGCTGGAGATGG 367	
Db	341	TGAGAGATTAATTGTGGAGGTAAAGAAAGATA---TTCAAATTCATGTACCTGGATGATGG 397	
Qy	368	AGG-----CAATGATCTCCTTTCAACGACTTTTGCTGCTCGATGACGTTGA 412	
Db	398	AAGATTAATGAGTTCCTTTAATGATCTCATGCAACGCCCTTTGGATAGTCGATAGCGTTGA 457	
Qy	413	ACGTTTCGGGAATCGACAGGCAATTTCAAAAAGAGATAAAAACGGCACTCGAATTATGTTAA 472	
Db	458	ACGTTTCGGGATAGCTAGACATTTCAAGAACGAGATAACATCAGCTCTGGATTATGTTTT 517	
Qy	473	CAGTTATTGAAACGAAAAAGGCATTGAGATGTGGGAGGAGAGTGTGTGACTGACCTCAA 532	
Db	518	CCGTTACTGGGAGAAACCGGCATTGGAATGTGGGAGAGACAGTATTGTTACTGATCTCAA 577	
Qy	533	CTCAACCGCTTCGGGCTTCGAACTCTCCGACTACACGGATACACTGTGCTTCAGATGT 592	
Db	578	CTCAACTGCGTGTGGGGTTTCGAACTCTTCGATTACACGGGTACACTGTATCTCCAGAGGT 637	
Qy	593	TTTGAACGTTTTTAAAGACAAAAATGGCAATTTTCTCCACTGCCAATATTTCAGATAGA 652	

Db 638 TTTAAAGCTTTTCAAGATCAAAATGGACAGTTTGTATGCTCCCCC---GGTCAGACAGA 694  
 Qy 653 GCGAGAGATTAGAGCGTCTCTCAATTTATTCAGGCGCTCCCTCGTCCCTTTCCCGCGA 712  
 Db 695 GGGTGAGATCAGAGCGTCTCTCAATTTATATCGGGCTTCCCTCATTTGCGCTTCCCTGGTGA 754  
 Qy 713 GAAAGTTATGATGAGCTGAAGCTGAACATTTCTCTCAAAAATATTAAAGAGAGCCCTGCAGAAA 772  
 Db 755 GAAAGTTATGAGAGAGCTGAATCTCTCCCAAGATATTGAAAGAGCTCTACAAA 814  
 Qy 773 GATTCCGGCAATCAGATATACTTTCAVAGAGATACGGAGCTTCTGGAATATGTTGGCA 832  
 Db 815 GATTCCAGTCTCCGCT---CTTTCAAGAGATAAAGTTTGTATGGAATATGCTGGCA 871  
 Qy 833 CACCAATTTCCAGCTTGAAGCAAGAGATTTACATGAGCTCTTTGGACAGACACTA- 891  
 Db 872 CACAAATTTGCCAAGATTGGAAGCAAGAAATTTACATAGACACACTTTGAGAAAGACACAG 931  
 Qy 892 -----AAAATAAGAACCCGCGAGAAACTTTTAGAACTTGCAGAAATTTGAAATCAA 943  
 Db 932 TGCATGGCTCAATAAAATGCTGGAGAGAGCTTTTAGAACTTGCAGAAATTTGAGTTCAA 991  
 Qy 944 TATATTTCACTCTTACAGAGAGAGAGTTAAACATGTTTCCGATGGTGGAGAAAGCTC 1003  
 Db 992 TATATTTAACTCTCTTACAAACAAAGGAAATTTACATATCTTTGAGATGGTGGAAAGAGTC 1051  
 Qy 1004 GGGTCTCTGAGATGACCTTCTGTGACATCGTACGCTGGAATACATGAGCTTTGGCTTC 1063  
 Db 1052 GGAATTTGCTTAATGACATTTGCTGGCAATGCTGATGGAATTTACACTTTGGCTTC 1111  
 Qy 1064 CTGCATTTGCTCGAGCTCAACATTTCTGGAATTCAGATCTGGCTTTTACCAAGATGTTCTCA 1123  
 Db 1112 TTGTATTTGCTTGAACCAAAACATTTCTGATTCAGACTAGGCTTCGCCAAATGTGTCA 1171  
 Qy 1124 TCTATCACGGTCTTTCAGACATGTCAGAGCTTCTCGGACAGATGACAGCTGGAAT 1183  
 Db 1172 TCTTGTACAGTTTTCGACATATTTTACGACACTTTTGGAAACGATTGACAGCTTGAAT 1231  
 Qy 1184 CTTTACAGCGACAAATTAAGAGATGGGATCGTCCGCGATGGAATGCTTCCAGAAATAT 1243  
 Db 1232 CTTTCAATCTGCAATTAAGAGATGGAATTCATCAGATAGAACACCTTCCAGATATAT 1291  
 Qy 1244 GAAAGAGTGTACATGATGTTTATCACACCGTAAATGAAATGGAATGCTCGAGTGGCAGAA 1303  
 Db 1292 GAAATGTGTACATGCTGTGTTTGAACCTGTAATGAACTGACACAGAGCGGAGAA 1351  
 Qy 1304 GGTCTAAGCGCGAGACAGCTCACTATGCAAGACAGCTTGGAGCGTGTGTTGATTC 1363  
 Db 1352 GACTCAAGGGAGAAACACTCTCACTATGTTTGAAGGGCTTGGAGGCTTATTTGATTC 1411  
 Qy 1364 GTATATGCGAAGCAAGTGAATCGCACTGGTTATCTGCCACGTTTTCAGAGTACTT 1423  
 Db 1412 ATATATGGAAGACAAATGGATCTCTAATGGTTATCTGCCAACGTTTGAAGATACCA 1471  
 Qy 1424 GGAAGCGGAAAGTTAGTCTGCTCATGCCCATGCGCATCTGCAACCCCATCTGAGTT 1483  
 Db 1472 TGAGATGGGAAAGTGAAGTCTGATATCGTAGCAACATTTGCAACCCATCTCACTTT 1531  
 Qy 1484 GGCATCCCTTTCTGATCAGCTCAAGAGTGAATGCTTCCCATCGAGCTCAATGA 1543  
 Db 1532 GAATGATGCTTCTGATTAATCTTGAAGGAAATTTGATTTTCCATCCAGGTTCAATGA 1591  
 Qy 1544 CTTGATATGATCATCTCTCGATTAAGAGGTGATACACGGTGTCTACAAGGCAGACAGGGC 1603  
 Db 1592 TTTGGCATCGTCTCTCTCGGCTACGAGGTGACACAGCTGTCTACAAGSCGATAGGA 1651  
 Qy 1604 CGGTGAGAGAGAGCTTCTGCTATATCATGTTTATGAAAGACAAATCTCGATTAACGA 1663  
 Db 1652 TCGTGTGAAGAGCTTCTGCTATATCATGTTTATGAAAGACAAATCTCGATTAACGA 1711  
 Qy 1664 AGAAGATGCTCTCAATCATATCAATTTATGATCAGGAGCGCAATCAGAGAAATTAATTC 1723  
 Db 1712 AGAAGATGCTCTCAATCATATCAATGCTGCTGATGATCAATCAAGAAATTAATTC 1771

Qy 1724 GGAGCTTCTAAAGCCAGACAAAGTGTTCCTCATCTTCCAGAAACACGCAATTTGACAT 1783  
 Db 1772 GGAATCTTGAAGTCAACGACATATTCATGCTGGCCAGAAACATGCTTTTGACAT 1831  
 Qy 1784 AAGCAGAGTTTGGCATCACGTTTACAGATACCGAGATGGCTACAGCTTTCCCAACGTTGA 1843  
 Db 1832 AACAGAGCTCTCCACCATCTCTACATATATCAGATGGCTTTAGTGTGGCCAAAGGA 1891  
 Qy 1844 AACAGAGTTTGGTATGAGAACCGTCATTTGAACCTTGCCCTTGTAAACACACTTCAA 1903  
 Db 1892 AACAAAAAATTTGGTTATGGAAACACTCTCTGAACTCTATGCTTTTAACTATAACCAT 1951  
 Qy 1904 ATCTACAATTAATAACTGAGGATCCCTATGCTGGTGTATATAGGCGACA 1950  
 Db 1952 TCCATATAATAGCTCATATGCTAAATTTATGGCTTATGACATA 1998

RESULT 11

AAF73391  
 ID AAF73391 standard; cDNA; 2205 BP.

AC AAF73391;

DT 30-APR-2001 (first entry)

Grand fir monoterpene synthase coding sequence fragment SEQ ID NO: 31.

Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;

myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;

terpinolene synthase; insect resistance; nutrition; ss.

Abies grandis.

WO200107565-A2.

01-FEB-2001.

24-JUL-2000; 2000WO-US020264.

26-JUL-1999; 99US-00360545.

(UNIW ) UNIV WASHINGTON STATE RES FOUND.

Steele CL, Bohlmann J, Croteau RB, Phillips MA;

WPI; 2001-182782/18.

P-PSDB; AAB69380.

New nucleic acid encoding monoterpene synthases, for increasing terpene synthesis in plants, e.g. for increasing resistance to pests or for treatment of cancer.

Example 3; Page 135-137; 175pp; English.

The present invention provides the protein and coding sequences of monoterpene synthases from the grand fir. These include (-)-camphene synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase and pinene synthase. The sequences can be used to produce transgenic plants expressing high levels of the enzymes, resulting in levels which are useful in protecting against and treating cancers, and to confer insect resistance on plants

Sequence 2205 BP; 703 A; 412 C; 463 G; 627 T; 0 U; 0 Other;

Query Match 53.2%; Score 1071.6; DB A; Length 2205;

Best Local Similarity 75.0%; Pred. No. 3.1e-293;

Matches 1430; Conservative 0; Mismatches 441; Indels 36; Gaps 6;

Qy 68 CAGTGTGCTGCTCAGTTCTTCTTCATGAGATTAAGGCTCTCCGTAGAACAAATCCCAACTCT 127

Db 104 CAAGTGTGATCAGTTCAATTCATGACATAGAGCTCCCTATAGAACAAATCCCAACTCT 163

128 TGGAACTCTGAGGCGGGGAAATCCGTCGGCATTCATCCATAAATGATGTTTGAACAGCGT 187  
164 TGGAACTCGGTAGCGGAGGAAATCTGTCAAGCTTCCATGACATAGTTTGGCCACCGC 223  
188 CGCATCTACTGATTTCTGTACAGAGACGGTGGGCAACTATGATTTCCAACTGTGGGACGA 247  
224 TGCACCTGATGATGGTTGTACAAAGACGCAATAGTGACTTACCATTTCCAAATATCTGGGACGA 283  
248 TGATTTTATACAGTCTCTGATCTCAAGCCCTTATGGAGCACTGATTTACCGGGAACGTGC 307  
284 TGATTTTATACAGTCTC---TATCAAGCATTTATGGGAACCCCTTACACAGGAACGTGC 340  
308 TGACAGACTTTATGGGAAGTAAAGGATATATGTTCAATTTCAAGTCGCTGGAAGATGG 367  
341 TGAGAGATTAATTTGGGGTAAAGAGATA---TTCAATTTCAATGTACCTGGATGATGG 397  
368 AGG-----CAATGATCTCTTCAACGACTTTTGTCTGGTGGTGGATGATGGA 412  
398 AAGATTAAATGATGTTCTTTAATGATCTCATGCAAGCCCTTTGGATAGTTCGATAGCTTGA 457  
413 ACGTTTGGGAATCGACAGCAATTTCAAAAGAGATTAAGGCGCACTCGATTTATGTTAA 472  
458 ACGTTTGGGGATGCTAGACATTTCAAGACGAGATAACATCAGCTCTGGATTAATGTTTT 517  
473 CAGTTATTGGAAACGAAAAAGGCAATGGATGTGGGAGGAGAGTGTGTGACTGACCTCAA 532  
518 CCGTTACTGGGAGAAACGCAATGGATGTGGGAGAGACAGTATTTGTTACTGATCTCAA 577  
533 CTCAAACCCCTTGGGGCTTGGAACTCTCCGACTACCGGATACATGTTCTTCAAGTGT 592  
578 CTCAACTCGTTGGGGTTTGGAACTCTTCGATTAACACGGGTACACTGTATCTCCAGAGGT 637  
593 TTTGAACGTTTAAAGACAAAATGGGCAATTTTCTCCACTGCGCAATATTCAGATAGA 652  
638 TTTAAAGCTTTCAAGATCAAAATGGACAGTTTGTATGCTCCCCC---GGTCAAGACAGA 694  
653 GGGAGAGATTAGAGCGGTTCTCAATTTATTCAGGGCTCCCTCGTCGCTTTCCCGGCGA 712  
695 GGGTGAATCAGAAGGTTCTTAACTTATATCGGGCTTCCCTCAATGCTTCCCTGGTGA 754  
713 GAAAGTTATGGAATGAAGCTGAAACATTTCTCAAAATATTTTAAAGAAAGCCCTGCAAAA 772  
755 GAAAGTTATGGAAGAGCTGAAATCTTCTCCAAAGATATTTGAAAGAGCTCTCAAAA 814  
773 GATTCGGCATCCAGTATATCTTCACTAGAGATACGGAGCTTCTGGAATATGGTTGGCA 832  
815 GATTCAGTCTCCGCT---CTTTTCAAGAGATAAAGTTTGTATGGAATATGGCTGGCA 871  
833 CACCAATTTGCCACGCTTGGAAAGCAAGGAATTAATGAGACGCTTTTGGACAGCACATA- 891  
872 CACAAATTTGCCAAGATTGGAAGCAGAAATTAATAGACACACTTTGAGAAAGACACAG 931  
892 -----AAAAAAGAACCGCGCCGAGAAACTTTTGAAGCTTGCAAAATTTGAATCAA 943  
932 TGCATGGCTCAATAAATAATGCTGGGAGAAAGCTTTTGAAGCTTGGCAAAATTTGAGTTCAA 991  
944 TATATTTCACTCTCTTCAAGAGAGAGAGTTTAAACATGTTTCCCGATGGTGAAGACTC 1003  
992 TATATTTAACTCTCTTCAACAAAGAAATTAATATCTTTTGAATGTTGGAAGAGTCT 1051  
1004 GGGTCTCTCTGATGATGACCTTCTGTCCGACATCTGTCAGTGGATACACTACGTTTGGCTTC 1063  
1052 GGAATTTGCCCTAAATTCAGATTTGCTCGGCATCTGATGTTGAATTTTACACTTTGGCTTC 1111  
1064 CTGCAATGGCTTGCAGCCTCAACATTTCTGATTTCACTCGGCTTACCAAGATGCTCA 1123  
1112 TTGATTTGCCATTTGACCCAAACATTTCTGATTTCACTAGGCTTCGCCAAATTTGTCTCA 1171  
1124 TCTTATCAGGTTCTTTCAGACATGATGACAGCTTCTCGGCACAGTACAGGCTGGAAT 1183  
1172 TCTTGTACAGATTTTGGACGATATTTACGACACTTTTGGAAAGATTTGACGAGCTTGAAT 1231

1184 CTTCAAGCCGACAAATTAAGATGGGATCCGTCGGATGGAATGCTTCCAGAAATATAT 1243  
1232 CTTTCAATCTGCAATTAAGAGATGGAAATCATCAGAGATAGAACACCTTCCAGATATAT 1291  
1244 GAAAGGAGTGTACATGATGTTTATCACCGTAAATGAAATGGCTCGAGTGGCAGAGAA 1303  
1292 GAAATGTGTATCATGTCGTGTTGAACTGTAAATGAACTGACAGAGAGCGGAGAA 1351  
1304 GGTTCAGGCGGAGACACGCTCAACTATGCAAGACAGCGCTTGGAGGGGTGTTTGAATC 1363  
1352 GACTCAAGGAGAGAAACACTCTCAACTATGTTTGAAGGCTTGGAGGCTTATTTTGAATC 1411  
1364 GTATATCGAGGAGCAAAAGTGGATCGGCACTGTATCTGCCACGTTTGAAGGAGTACTT 1423  
1412 ATATATGGAAGAGCAAAATGGATCTCTAATGGTTATCTGCCAACGTTTGAAGAGTACCA 1471  
1424 GGAAACGGGAGAACTTAGCTCTGCTCATCGCCCATGCGCACTGCAACCCATTTCTGAAGCTT 1483  
1472 TGAGATGGGAAAGTGGAGCTCTGCATATCGGTAGCAACATTTGCAACCCATCTCCTCCTT 1531  
1484 GGACATCCCTTTCTGATACATCTCAAGGAAAGTGTGACTTCCATCGAAGCTCAATGA 1543  
1532 GAATGCAATGGCTTCTGATTAACATCTTGAAGGAAATGATTTTCCATCCAGGTTCAATGA 1591  
1544 CTTGATATGATCATCTTCTGATTAAGAGGTATACCGGTGCTACAGGCGAGACAGGCG 1603  
1592 TTTGGCATCGCTTCTTCTTCTGAGGTGACACGCTGCTACAGGCGGATAGGGA 1651  
1604 CCGTGGAGAAAGCTTCTGCTATATCATGTTTATGAAAGCAATCTCTGGATTAACCGA 1663  
1652 TCGTGGTGAAGAGCTTCTGATATCATGTTATGAAAGCAATCTCTGGATCAACCGA 1711  
1664 AGAAGATCTCTGAATCATATCACTTCAATGATCAGGACCGAATCAGAGATTAATG 1723  
1712 AGAAGATCCCTCAATCATATCAATGCTTCAATGATGATTAATCAAGAAATTAATG 1771  
1724 GGAGCTTCTTAAGCCAGACACAGCTTCCCATCATCTTCCAAAGAAACGCAATTTGAAT 1783  
1772 GGAATCTTAAAGATCCAACGCAATATTTCCAAATGCTGGCCAAAGAAACATGCTTTTGACAT 1831  
1784 AAGCAGAGTTTGGCATCACGGTTACAGATACCGAGATGCTACAGCTTTGCCAAGCTTGA 1843  
1832 AACAGAGCTCTCCACCATCTCTACATATATCGAGATGCTTTAGTGTGCCAAGAGGA 1891  
1844 AACAAAGAGTTTGGTGAAGAACCGCTCATGGAACCTGTCCCTTTGTAACACACTTCAA 1903  
1892 AACAAAAAAATTTGGTTATGGAACACATCTCTTGAATCTATGCTTTTAACTATAACCCATA 1951  
1904 ATCTACATATTAATCTGAGGATGCCCTATGGGTGATATATAGGGCACA 1950  
1952 TCCATAATAATAAGCTCATATGCTAAATTTATTTGGCTTTATGACATA 1998

RESULT 12  
AAE73413  
ID AAE73413 standard; cDNA; 2429 BP.  
XX  
AC AAE73413;  
XX  
DT 30-APR-2001 (first entry)  
XX  
DE Grand fir monoterpene synthase coding sequence fragment SEQ ID NO: 68.  
XX  
KW Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
KW terpinolene synthase; insect resistance; nutrition; ss.  
XX  
OS Abies grandis.  
XX  
PN WO200107565-A2.  
XX  
PD 01-FEB-2001.

PF 24-JUL-2000; 2000WO-US020264.  
 PR 26-JUL-1999; 99US-00360545.  
 XX (UNIW) UNIV WASHINGTON STATE RES FOUND.  
 PA Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
 PI WPI; 2001-182782/18.  
 XX DR P-PSDB; AAB69392.  
 PT New nucleic acid encoding monoterpene synthases, for increasing terpene  
 PT synthesis in plants, e.g. for increasing resistance to pests or for  
 PT treatment of cancer.  
 XX Claim 23; Page 156-159; 175pp; English.  
 XX The present invention provides the protein and coding sequences of  
 CC monoterpene synthases from the grand fir. These include (-)-camphene  
 CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 CC limonene, (-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 CC and pinene synthase. The sequences can be used to produce transgenic  
 CC plants expressing high levels of the enzymes, resulting in levels which  
 CC are useful in protecting against and treating cancers, and to confer  
 CC insect resistance on plants  
 XX SQ Sequence 2429 BP; 726 A; 484 C; 512 G; 707 T; 0 U; 0 Other;

Query Match 46.2%; Score 929.4; DB 4; Length 2429;  
 Best Local Similarity 71.1%; Pred. No. 7.7e-253;  
 Matches 1335; Conservative 0; Mismatches 491; Indels 51; Gaps 6;

QY 72 TCCTGCTCAGTCTTCTCTCATGAGATTAAAGGCTCTCCGTAGAACAAATCCCAACTCTTGGA 131  
 DB 89 TCCTTGATCAGTCTCCAGCAATGTGCAGAGGCTCTCTGTATCTCTACAGAGTCCCAACT 148  
 QY 132 ATCTGAGCGCGGGAAATCCGTCGGCATTCATTAACAATGTGTTTGAACAAGTCGCA 191  
 DB 149 CTGAGAAATGGTATGAGGAGACAAAGGCTCTGTCTCATCAACATGAATGACCACTGATCC 208  
 QY 192 TCTACTGATCT 236  
 DB 209 CATCTGATGATTAATGGT 268  
 QY 237 CTGTGGAGAGATGATTTTCAATGATCTCTGATCTCAACGCTTATGGAGCACCTGATTA 296  
 DB 269 CTGTGGAGAGATGATTTTCAATGATCTCTGATCTCAACGCTTATGGAGCACCTGATTA 325  
 QY 297 CGGGAACGCTGACAGACTTATGGGAGTAAAGGATATATATATATATATATATATATATAT 347  
 DB 326 AGTGAACGCTGTGACAGTGTGTGAGGAGTAAAGGATGTTCAATTAATACCAAT 385  
 QY 348 TTCAAGTCGCTGGAAGATGAGGCAATGATCTCTTCAACGACTTTTGTGTCGATGAC 407  
 DB 386 AATAGAGATTAATTTGGTTCCTCAATGATCTCTTACAGCCTTTGGATGGTGGATAGC 445  
 QY 408 GTTGAACGCTTGGGATGACAGGCAATTTCAAAAAGAGATAAAAACGGCACTCGATTAT 467  
 DB 446 ATTGAACGCTTGGGATGACAGCAATTTCCAAAATGAGATAAGAGTAGTCCCTCGATTAT 505  
 QY 468 GTTAACAGTTATTGGAACGAAA- --AAGCAATTTGGATGGGAGGAGAGTGTGTGACT 524  
 DB 506 GTTTACAGTTATTGGAAGAAAGGAAGGCAATTTGGGTGGGAGAGATTTCTACTTTTCCT 565  
 QY 525 GACCTCAACTCAACGCTTTGGGGCTTCGAACTCTCCGACTACACGATACACTGTGTCT 584  
 DB 566 GATCTCAACTGACTGCTCTGGCGCTTCGAACTCTTCGACTGCAAGGATACAAATGTGTCT 625  
 QY 585 TCAGATGTTTGAACGTTTTHAAGCAAAAATGGGCAATTTTCTTCACTCCCAATATT 644  
 DB 626 TCAGATGTTGGAATACTTCAAGATCAAAAGGGGCAATTTTGCCTGCGCTGCAATCTCTA 685  
 QY 645 CAGATAGAGGAGAGATTAGAGGCGTCTCTCAATTTATTACAGGCGCTCCCTCTCGCTTT 704

DB 686 ACCGAGGACAGATCCTAGAGTGTCTAAATTTATATCGGGCTTCCCTGGTCCCTTT 745  
 QY 705 CCCGCGAGAAAGTTATGGATGAAGCTGAACATTTCTCTCAAAATATTTTAAGAGAGCC 764  
 DB 746 CCGGGGAGAAAGTTATGAAGAGGCTGAATCTTCTCGGCATCTTATTTGAAAGAGTC 805  
 QY 765 CTGCAAAAGATTCCCGCATCCAGTATCTTCTACATAGATACGGAGCTTCTGGAATAT 824  
 DB 806 TTACAAAGATTCCAGTCTCCAGT- --TTTTCAGAGAGATAGAAATACGTTTGGGAATAT 862  
 QY 825 GGTGGCACCAATTTCCACGCTTGAAGCAAGAGAAATACATGGAGCTCTTTGGAGCAG 884  
 DB 863 GGTGGCACCAATTTCCCAAGATTGGAAGCAAGAAATATATATCGAGCTCTACGGGAG 922  
 QY 885 CACTATAAATAAGAACGCG- --CCGAGAACTTTTAAACTT 926  
 DB 923 GACAGTATGAAGTTCAAAACGAGATGCCATATGTGAATAGCGAAGCTTTTAAACTT 982  
 QY 927 GCAAAATTGGAATTCATATATATTTCACTCTTACAGAGAGAGATTTAAACATGTTTCC 986  
 DB 983 GCAAAATTGGAATTCATATATATTTCACTCTTTCACAGAGAAAGATTTGCAATATATCT 1042  
 QY 987 CGATGTTGAAAGACTCGGCTTCTCTGAGATGACCTTCTGTGACATCTGTCACTGCGAA 1046  
 DB 1043 AGATGTTGAAAGATTCTGTTCTCATCTCATCTGACTTTTACTCGACACCTGACGTGAA 1102  
 QY 1047 TACTACGCTTTGCTTCTCCTGANTGCTTCCAGCTTCAACATTTCTGAGATCAGACTCGG 1106  
 DB 1103 TACTACCAATGGCATCTTCAATTTCTATGGAGCGAAGACACCTCGCTTTTCAAGTTGGG 1162  
 QY 1107 TTTTACCAAGATGCTCATCTTATCAGGTTCTTGAAGACATGTACGAGCTTCTTGGGACA 1166  
 DB 1163 TTTGTCAAAACATGCTCATCTTCTAACAGTTCTCGATGATGATGATGACACTTTTGAACA 1222  
 QY 1167 GTACAGAGCTGGAATCTTTTACAGGCAAAATTAAGAGATGGATCGTCCGCGATGAA 1226  
 DB 1223 CTGAGCACTCCCACTTTTACGCTGCTTTAAGAGATGGATTTGTGAGAGACAAAG 1282  
 QY 1227 TGCTTCCAGATATATGAAGAGGTGATCATGATGGTTTATCACCGTAAATGAATG 1286  
 DB 1283 TGTCTTCCAGATATATGAAGCAGTGTACATGCACTTGTATCAATGTCTTAATGAATG 1342  
 QY 1287 GCTCGAGTGGCAGAGAGCTCAAGCGCAGACACGCTCAACTCAACTGCAAGAGCTTGG 1346  
 DB 1343 GCGCAAGAGCTGAGAGACTCAAGCAGAGATACGCTCAACTATATATTCGAATGCTTAT 1402  
 QY 1347 GAGCGTGTGTTGATTCGATATGAGAGGAGCAAAAGTGGATCGCACTGTTTCTGCCC 1406  
 DB 1403 GAGTCTCAATTTGATTCGTTTATGCAAGCAAAATGGATCTCAAGTGTATCTCCCA 1462  
 QY 1407 ACGTTTGAAGTACTTGGAGAACGGGAAAGTTAGTCTGTCTCATCGCCCATCGGCATG 1466  
 DB 1463 ACGTTTGAAGTACTTGAAGAAAGTGGTTCGGTTCTCGCAGCACTTTA 1522  
 QY 1467 CAACCAATTCACGTTGGACATCCCTTTTCTGTATCATCTCTCAAGAGTTGACTTC 1526  
 DB 1523 CAACCAATTCACCTTGGATGTACCACTTCTTAATTAATTAATTAATTAATTAATTAAT 1582  
 QY 1527 CATCGAAGCTCAATGACTTGAATGTATCATCTTCTGATTAAAGGATGATACCGGTGC 1586  
 DB 1583 CCATCTAGTTCATGACTTGGCTTGGTCCCTTCTCGCTACGTGGTGCACGCGTGC 1642  
 QY 1587 TACRAGGAGACAGCGCGCTGGAGAGAGCTTCTGTATCATCTGTTATTAAGAGAC 1646  
 DB 1643 TACRAGGAGATGAGGCTGGGAGAGAGCTTCAAGTATATCTGTTTATTAAGAGAC 1702  
 QY 1647 AATCTGGAATTAACGGAAGAGATCTCTGAATCATATCAACTTCAATGATCAGGAGCA 1706  
 DB 1703 CATCTGATCAACAGAGAGAGATGCTCTCAATCATATCAAGCTCATGATCATGATGCA 1762  
 QY 1707 ATCAGAGATTAATTTGGAGCTTTAAAGCGCAGCAACAGTGTGTTCCCATCACTTCCAAG 1766

Db 1763 ATCAGAGAAATTAATTGGGAGCTTCTCAGACCAGATAGCAAAAGTCCCAATCTCTTCCAAG 1822  
QY 1767 AAACAGCATTTTACATAAGCAGATTTGGCATCAGCGTTTACAGATACCGAGATGGCTAC 1826  
Db 1823 AAACATGCTTTTACATCACCAGAGCTTCCATCACCTCTACAAGTACCGAGATGGTTAC 1882  
QY 1827 AGCTTTGCCACCTTTGAACAAGAGTTTGGTGATGAGAACCGTCAATTGAACCTGTGGCT 1886  
Db 1883 ACTGTTCGAGTAGTGAACAAGAAATTTGGTGATGAACAACAGTTCTTGAACCTGTGGCA 1942  
QY 1887 TTGTACACACTTCAA 1903  
Db 1943 TTGTAAAAAATAATCAAA 1959

## RESULT 13

AA08645  
ID AAX08645 standard; cDNA; 2089 BP.

AC AAX08645;

XX 27-SEP-1999 (first entry)

DE Limonene synthase gene.

XX Myrcene synthase; limonene synthase; pinene synthase; pinene synthase; flavour; aroma;  
KW defense; plant seed; oil; meal; ss.

XX Abies grandis.

OS Location/Qualifiers  
FH 73..1986  
FT /\*tag= a  
FT /product= "Limonene synthase"

XX WO9902030-A1.

XX 21-JAN-1999.

XX 10-JUL-1998; 98WO-US014528.

XX 11-JUL-1997; 97US-0052249P.

XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX Bohlmann J, Steele CL, Croteau RB;

XX WPI; 1999-120396/10.

XX P-PSDB; AAW85702.

XX New isolated gymnosperm monoterpene synthase DNA - obtained from Grand  
PT fir (Abies grandis), used to provide plants with modified production of  
PT monoterpenes, e.g. myrcene, limonene or pinene.

XX Claim 11; Page 79-82; 121pp; English.

XX Nucleotide sequences encoding myrcene synthase, limonene synthase and  
CC pinene synthase from Grand fir may be incorporated into any organism  
CC (e.g. intact plant, animal, microbe), or derived cell culture that  
CC produces geranyl diphosphate for the production of the aforementioned  
CC enzymes or their products. The sequences when expressed in transfected  
CC cells may also be used for the production or modification of flavour and  
CC aroma properties, improvement of defense capability, and the alteration  
CC of other ecological interactions mediated by myrcene, limonene, pinene,  
CC or their derivatives. In particular they can be used for the production  
CC of plant seeds for the extraction of oil or meal

XX Sequence 2089 BP; 638 A; 415 C; 450 G; 586 T; 0 U; 0 Other;

Query Match 46.0%; Score 925.4; DB 2; Length 2089;

Best Local Similarity 70.6%; Pred. No. 9.8e-252;

Matches 1340; Conservative 0; Mismatches 506; Indels 51; Gaps 6;

QY 72 TGTGCTCAGTTCTTCTCATGAGATTAAGGCTCTCCGTAGAACAAATCCCAACTCTTGGGA 131  
Db 127 TCGTTGATCAGTTCCAGCAATGTGCAGAAAGGCTCTCTGTATCTCTACAGCAGTCCCAACA 186  
QY 132 ATCTGAGCCCGGGGAATCCGTCGGCATTCCATAAATCATGTGTTTGACAAGCGTCGCA 191  
Db 187 CTCAGAAATCGTAGGGCAGAGAAAGCTCTGGTTCATCAATGAAATGACCATCTGATCC 246  
QY 192 TCTACTGATTCT-----GTACAGAGACGGGTGGCAACTATCATTCACAAC 236  
Db 247 CATCGTGATCAATATCGTGTGTCTGCTGCAAGACGCATAGCCGATCATCATCCCAAC 306  
QY 237 CTGTGGGACGATGATTTTATACAGTCTCTGATCTCAACGCTTTATGGAGACCTCGATTAC 296  
Db 307 CTGTGGGAAGATGATTTTATACAAATCATTTGCTCTCA---CCTTATGGGGATCTTGGTAC 363  
QY 297 CGGGAACGTCTGACAGACTTATTTGGGAAGTAAGGATAT-----AAGTTCAAT 347  
Db 364 AGTGAACGTCTGACAGCTGCTTCCCAAAATGATCTCTTACAGCCTTTGGATGGGTAGC 423  
QY 348 TTCAGTCTGTGAGATGAGGCAATGATCTCTTCAACGACTTTTGTGTCGATGAC 407  
Db 424 AATAGAGAATTAATTTGGTTCCCAAAATGATCTCTTACAGCCTTTGGATGGGTAGC 483  
QY 408 GTTGAACGTTTGGGAATCGACAGGCAATTTCAAAAAGAGATAAAAACGGCACTCGATTAT 467  
Db 484 ATTGAACGTCTGGGGATAGATAGACATTTCCAAAATGAGATAGAGTAGCCCTCGATTAT 543  
QY 468 GTTAACAGTATTGGGAACGAAA---AAGGCAATTTGGATGTTGGAGGAGAGTGTGTGACT 524  
Db 544 GTTTACAGTTATTTGGAAGGAAGAGGCAATTTGGGTGTGGCAGAGATTTCTATTTCCT 603  
QY 525 GACCTCAACTCAACCGCTTTGGGGCTTCGAACCTCTCCGACTCAGCAACGATACACTGTGTCT 584  
Db 604 GATCTCAACTCGACTGCTTGGCGCTTCGAACCTCTTCGACTCTTCAGCGATACAAATGTCT 663  
QY 585 TCAGATGTTTGAACGTTTAAAGACAAAATGGGCAATTTTCTCCCACTGCAATATTT 644  
Db 664 TCAGATGCTGTGGAATACTTCAAGATGAAAGGGGCAATTTTCTGCTGCTTCAATCTTA 723  
QY 645 CAGATAGAGGAGAGATTAAGGCGCTTCAATTTATTCAGGCGCTTCCCTCGCTTTT 704  
Db 724 ACCGAGGACAGATCACTAGAGTGTCTAAATTTATTCGGGCTTCCCTTGGTGGCTTT 783  
QY 705 CCGGCGAGAAAGTTATGGATGAAGCTGAAACATTTCTTACAAAATATTTAAGAGAAGCC 764  
Db 784 CCGGGGAGAAAGTTATGGAAGAGGCTGAAATCTTCTCGGCACTCTTATTTGAAAAGTCT 843  
QY 765 CTGCAAAAGATTCGGCATCCAGTACTTTTCACTAGAGATACGGGACGTTCTGGAATAT 824  
Db 844 TTACAAAAGATTCGGCTCTCCAAT---CTTTCAGGAGAGATAGAAATATGTTTGGAAATAT 900  
QY 825 GGTGTGCACACCAATTTGCCACGCTTTGGAAGCAAGAAATTTACATGGAAGCTTTTGGACAG 884  
Db 901 GGTGTGCACACGAATTTGGCGAGATTTGGAAGCAAGAAATTTATATGAGGTCTACGAGCAG 960  
QY 885 CACACTAAAAATA-----AGAACCGCCCGAGAAACCTTTTAGAACTT 926  
Db 961 AGCGGCTATGAAGCTTAAACGAGATGCCATATATGAACATGAAGAAGCTTTTACAACCT 1020  
QY 927 GCNAAATTTGAATTTCAATATATTTTCACTCTTACAGAGAGAGAGTTAAACATGTTTCC 986  
Db 1021 GCNAAATTTGAATTTCAATATATTTTCACTCTTTCGCAACTAAGAGAGTTCAATCTATCTCC 1080  
QY 987 CGATGTGTGAAGAGACTTCGGGTTCTCTCGTGAATGACCTTCTGTGACATCGTCAAGTGGAA 1046  
Db 1081 AGATGTGTGAAGAAATCAGGTTCTCTCACTGACTTTTACAGCGCATGCTACGTTGGAA 1140  
QY 1047 TACTACGCTTTGGCTTCTCGATTTGCGTTCGAGCCTCAACATTTCTGATTCAGACTCGGC 1106  
Db 1141 TACTACATATGCACTTTGCAATTTCTATGTTGCCAAAACATTTCAAGCTTTTCAAGATGGAG 1200  
QY 1107 TTTACCAAGATGCTCATCTTATCATCGGTTCTTGACGACATGTACGAGCTTCTTCGGCACA 1166

Db 1201 TTTGTCAAAGTGTCTCTTTTAAACAGTTCCTGATGATATATATGACACATTTTGGAAACA 1260  
 Qy 1167 GTACAGAGCTGGAACTCTTACAGCGACAATTAAGAGATGGATCCGTCGGATGAA 1226  
 Db 1261 ATGAACGAACCTCAACTTTTACGGATGCAATTAAGAGATGGATTTGTCACACGAAG 1320  
 Qy 1227 TGCCTTCCAGAAATATGAAGAGGTGTACATGATGTTTATCACACGTAATGAAGATG 1286  
 Db 1321 TGGCTTCCAGAAATATGAAGAGGTGTACATGATGTTTATCACACGTAATGAAGATG 1380  
 Qy 1287 GCTCGAGTGGCAGAAAGCTCAAGCGCGAGACACGCTCAACTATGTCACAAAGACAGCTTGG 1346  
 Db 1381 GTGAAGAGAGCTGAGAGAGCTCAAGCGCGAGATGCTCAACTATGTCACAAAGATGCTTGG 1440  
 Qy 1347 GAGCGGTGTTGATTCGTATATGACGAAGCAAGTGTGATGCGCAGCTGTTATGCGC 1406  
 Db 1441 GAAGCCCTATTGATACCTTTATGCAAGCAAGTGTGATGCGCAGCTGTTATGCGC 1500  
 Qy 1407 ACCTTGGAGAGTACTTGGAGAACGGGAAAGTTAGCTCTGCTCATGCGCCATGCGCACTG 1466  
 Db 1501 ACCTTGGAGAGTACTTGAAGATGCAAGAGTTAGTCTGCTGCTGATGATGATGATG 1560  
 Qy 1467 CAACCCATCTGAGCTTGGACATCCCTTTCTGATCATCTCTCAAGCAAGTGTGATGCT 1526  
 Db 1561 CAACCCATCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620  
 Qy 1527 CCATCGAGCTCAATGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1586  
 Db 1621 CCATCGAGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 Qy 1587 TACAAGCAGACAGCGCGCGGAGAGAGAGCTTCTGCTATATATATATATATATGAAAG 1646  
 Db 1681 TACAAGCAGAGTAGGCGCGGAGAGAGAGCTTCTGCTATATATATATATATGAAAG 1740  
 Qy 1647 AATCTGATTAACGGAAGAGATGCTCTGATCATATATATATATATATATATATGAAAG 1706  
 Db 1741 CATCTGATCAATGAGAGAGAGTCTCTGATCATATATATATATATATATATGAAAG 1800  
 Qy 1707 ATCAGAGATTAATGAGAGAGTCTTAAAGCAGACAGAGTGTGCTCATCATCTTCCAT 1766  
 Db 1801 ATCAGAGATTAATGAGAGAGTCTTAAAGCAGACAGAGTGTGCTCATCATCTTCCAT 1860  
 Qy 1767 AAACACGATTTGACATAAGCAGAGTGTGCTCATCATCTTCCATCATCATCTTCCAT 1826  
 Db 1861 AAACATGCTTTGATCATCAACAGAGTGTGCTCATCATCTTCCATCATCATCTTCCAT 1920  
 Qy 1827 AGCTTTGCCACGTTGAAACAAAGAGTGTGCTCATCATCTTCCATCATCATCTTCCAT 1886  
 Db 1921 ACTGTTTCCAAACAAAGAGTGTGCTCATCATCTTCCATCATCATCTTCCATCTCGCT 1980  
 Qy 1887 TTGTAACAAACACTTCAATCTCAATATTAATCTGAGG 1923  
 Db 1981 TTGTAACAAACATATAGATGATGATGATGATGATGATGATGATGATGATGATGATG

RESULT 14

AAA38938 standard; DNA; 2089 BP.

AC AAA38938;

XX 25-AUG-2000 (first entry)

DE Grand fir limonene synthase DNA sequence SEQ ID NO:57.

Synthase; protein co-ordinate data; active site; modification; terpenoid;  
 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;  
 isoprenoid; breeding programme; fragrance; flavour; pheromone;  
 defensive agent; pigment; antitumour; steroid hormone;  
 signal transduction pathway; bile acid; affinity purification;  
 photoreceptor; enzymatic synthesis; nutrient supplement;  
 immunological reagent; ds.

XX Abies grandis.  
 OS WO200017327-A2.  
 PN 30-MAR-2000.  
 PD 17-SEP-1999; 99WO-US021419.  
 PF 18-SEP-1998; 98US-0100993P.  
 PR 22-APR-1999; 99US-0130628P.  
 PR 23-AUG-1999; 99US-0150262P.  
 XX (KENT) UNIV KENTUCKY RES DEPT.  
 PA (SALK) SALK INST BIOLOGICAL STUDIES.  
 XX Chappell J, Manna KR, Noel JP, Starks CM;  
 PI WPI; 2000-292839/25.  
 DR P-PSDB; AAY90859.  
 XX Novel terpene synthase enzymes, useful for producing terpene  
 PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known  
 PT enzymes by specific amino acid alterations.  
 XX Disclosure; Page 445-448; 450pp; English.  
 XX The present invention describes an isolated terpene synthase (I)  
 CC comprising a region with at least 20% identity to region 265-535 of a 548  
 CC amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha  
 CC -carbon atoms (alphaC) that have interatomic distances, between each  
 CC other, within tabulated ranges, have a centre point (within a sphere of  
 CC radius 2.3 Angstrom) within tabulated ranges, and have an ordered  
 CC tabulated arrangements (tables given in the specification). (I), and  
 CC related enzymes, are used to produce a wide range of terpenoids (e.g.  
 CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,  
 CC flavours, pheromones, defensive agents, pigments, antitumour agents,  
 CC components of signal transduction pathways, precursors of steroid  
 CC hormones and bile acids, as photoreceptors and as co-factor side chains.  
 CC Some synthases with little or no catalytic activity (and nucleic acids  
 CC encoding them) are used as controls in the analysis of products formed by  
 CC enzymatic synthesis; as nutrient supplements; for affinity purification  
 CC of isoprenoids; or to develop immunological reagents or nucleic acids for  
 CC monitoring expression of terpene synthase or inheritance of the gene in  
 CC plant breeding programs. The new synthases may produce novel terpene  
 CC products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent  
 CC sequences used in the exemplification of the present invention  
 XX SQ Sequence 2089 BP; 638 A; 415 C; 450 G; 586 T; 0 U; 0 Other;

Query Match 46.0%; Score 925.4; DB 3; Length 2089;  
 Best Local Similarity 70.6%; Pred. No. 9.8e-252;  
 Matches 1340; Conservative 0; Mismatches 506; Indels 51; Gaps 6;

Qy 72 TCGTSCCTCAGTTCTTCTCATGAGATTAAGCTCTCCGTAGACAATCCCAACTTTGGA 131  
 Db 127 TCGTTGATCAGTTCAGCAATGTGCAGAGGCTCTCTGTATCTCTACAGCAGTCCCAACA 186  
 Qy 132 ATCTGAGGCGGGGAAATCGTCGCGCATTCCTAATAACATGTTTGTGACAGGCTGCA 191  
 Db 187 CTCAGAAATGCGTAGGCGACAGAAAGCTCTGTGTCATCAACATGAATTAACCACTGTATCC 246  
 Qy 192 TCTACTGATCTT-----GTACAGAGAGCGGTGGGCAACTATCATTCACAAC 236  
 Db 247 CATCGTATGATATGTTGGTGGTGTACTGCAAGACGCATAGCCGATCATCTCCCAAC 306  
 Qy 237 CTGTGGACGATGATTTCAACAGTCTCTGATCTCAACGCTTATGAGACACCTGATTAAC 296  
 Db 307 CTGTGGGAAGATGATTTTCAACATCATCTTGTCTCTCA---CCTTATGGGGATCTTCGTAC 363  
 Qy 297 CGGAAACGTCTGACAGACTTATTGGGGAAGTAAAGGATAT-----AATGTTCAAT 347

Db 364 AGTGAACGTGCTGAGACAGTCTGTTGAGGAAGTAAAGAGATGTTCAATTCAATACCAAAAT 423  
QY 348 TTCAAGTCGCTGGAAGATGGAGCAATGATCTCCCTTCAACGCACTTTTGTGCTGCGATGAC 407  
Db 424 AATAGAGAATATTGTTGTTCCCAAAATGATCTCTTACACGCTTTGGATGGTATAGC 483  
QY 408 GTTGAACGTTGGGAATCGACAGGCAATTTCAAAAAGAGATAAAAACGGCACTCGATTAT 467  
Db 484 ATTGAACGTCCTGGGATAGATAGACATTTTCCAAAATGAGATAAGAGTAGTCCCTCGATTAT 543  
QY 468 GTTAACAGTTATTGGAACGAAA- --AAGGCAATTGGATGTGGAGGGAGAGTGTGTGACT 524  
Db 544 GTTTACAGTTATTGGAAGAAAAGAGGCAATTTGGTGTGGCAGAGATTCTACTTTTCCT 603  
QY 525 GACCTCAACTAACCGCTTTGGGGCTTCGAACTCTCCCGACTACAGGATPACACTGTGTCT 584  
Db 604 GATCTCAACTCGACTGCTTGGCGTCTCGAACTCTTCGACTCGACGGATACAAATGTGTCT 663  
QY 585 TCAGATGTTTTGAACGTTTTTAAAGACAAAATGGGCAATTTTCTCCACTGCCAATATT 644  
Db 664 TCAGATGTGCTGGAATACCTTCAAAATGAAAAGGGGCATTTTGCCTGCCCTGCAATCTTA 723  
QY 645 CAGATAGAGGGAGAGATTAGAGCGTTCTCAATTTATTTCAGGGCCTCCCTCGTCCGCTTT 704  
Db 724 ACCGAGGACAGATCATTAGAAAGTGTCTAAATTTATTATCGGGCTTTCCCTGGTCCGCTTT 783  
QY 705 CCCGGCGAAGAAAGTTATGGATGAAGCTGAAACATTTCTTACAAAATATTAAAGAGAAAGCC 764  
Db 784 CCCGGCGAAGAAAGTTATGAAGAGGCTGAAATCTTCTCGGCACTCTTATTTGAAAAGAGTC 843  
QY 765 CTGCAAAAGATTCCGGCATCCAGTATACHTTCACTAGAGATAGGAGCTTCTGGAATAT 824  
Db 844 TTACAAAAGATTCCGGTCTCCAAAT---CTTTCCAGGAGAGATAGAATATGTTTTGGAATAT 900  
QY 825 GGTGGCCACCAATTTGGCAGCTTGGAAAGCAAGCAATTACATGACGCTCTTTGGACAG 884  
Db 901 GGTGGCCACCAATTTGCCGAGATTGGNAGCAAGAAATTATTCGAGTCTACAGCAG 960  
QY 885 CACATAAAATA- -----AGAACGCCGCCGAGAAACTTTTGAACATT 926  
Db 961 AGCGGCTATGAAGCTTAAACGAGATGCCATATATGAACATGAAGAAGCTTTTACAACATT 1020  
QY 927 GCAAAATTTGGAATTCATATATTTCACCTCTACAGAGAGAGACTTAAACATGTTTCC 986  
Db 1021 GCAAAATTTGGAGTTCAATATCTTTCACTCTTTGCAACTAAAGAGATTACAATCTATCTCC 1080  
QY 987 CGATGGTGAAGAAGACTCGGGTTCTCTCGAGATGACCTTCTGTCGACATCGTCACGTTGAA 1046  
Db 1081 AGATGGTGAAGAAGATCAGGTTGCTGCTCACTGACTTTTACACGGCACTGTCACGTTGAA 1140  
QY 1047 TACTACGCTTTGGCTTCTGCTGATTTGGTTTCGAGCCTCAACATCTTGGAATTCAGACTGGC 1106  
Db 1141 TACTACACTATGGCATCTTGCAATTTCTATGTTGCCAAAACATTCAGCTTTCAGAATGGAG 1200  
QY 1107 TTTACCAAGATCTCATCTTATTCAGGTTCTTGCGACATCTAGAGCTTTCGGCACA 1166  
Db 1201 TTTGTCAAAGTGTGTCATCTTGTAACAGTTCTCGATGATATATATGACACTTTTGAACA 1260  
QY 1167 GTPAGACGAGCTGGAACCTTTTCACAGCGACAATTAAGAGATGGGATCCGTCGCCGATGGAA 1226  
Db 1261 ATGAACGAACTTCAACTTTTACGGATGCAATTAAGAGATGGGATTTGTCAACGACAAGG 1320  
QY 1227 TGCCTTCAGAATATATGAAGAGTGTATACATGATGTTTATCACACCGTAATGAATG 1286  
Db 1321 TGGCTTCCAGATAATATGAAGAGGTGTATACATGGACTTGTATCAATGCAATTAATGAATG 1380  
QY 1287 GCTCGAGTGGCAGAGAGGCTCAAGCCGAGACAGCTCAACTATATCAAGACAGCTTGG 1346  
Db 1381 GTGGAAGAGGCTGAGAAGACTCAAGCCGAGATATGCTCAACTATATTCAAAATGCTTGG 1440  
QY 1347 GAGGCGTGTGTTGATTTCGTATATGCAAGGAAGCAAAAGTGGATCGCACTGGTTATCTGCC 1406  
Db 1441 GAAAGCCATTTTGATACCTTTATGCAAGAGCAAGTGGATCTCCAGCAGTTATCTCCA 1500

QY 1407 ACCTTTGAGGAGTACTTGGAGAACGGGAAGTTAGTCTGTCTCATCGCCCATGCGCACTG 1466  
Db 1501 ACGTTTGAGGAGTACTTGAAGAATGCAAAAGTTAGTTCTGTGGTTCTCGCATGCCACATTA 1560  
QY 1467 CAACCCATTCTGACGTTGGACATCCCTTTTCTGTATCAGATCCTCAAGGAAATTGACTTC 1526  
Db 1561 CAACCCATTCTCAGTTTGGATGTACCACTTCTCTGATTACATCTGCTCAAGAAATTGATTAT 1620  
QY 1527 CCATCGAAAGCTCAATGACTTTGATATGATTCATCCCTTCGATTAAAGAGGTGATACAGGTGC 1586  
Db 1621 CCATCCAGATTCAATGAGTTAGTCTTGGTCCATCTTCGACTACAGGTGACAGCGCTGC 1680  
QY 1587 TACAAGGCAGACAGGCCCCGTGGAGAAGAGCTTCGTCTATATCATGTGTTATATGAAAGAC 1646  
Db 1681 TACAAGGCGGATAGGCCCCGTGGAGAAGAGCTTCAGCTATATATCGTGTATTATGAAAGAC 1740  
QY 1647 AATCTCGAATTAAACGGAAGAGATGCTCTGAATCATATCAACTTCATGATCAGGGACGCA 1706  
Db 1741 CATCTCGATCAATAGAGAGAGATGCTCTCAATCATATCAACGCCATCATCAGTGATGCA 1800  
QY 1707 ATCAGAGAAATTAATTTGGAGCTTCTAAAGCCAGACAACAGGTGTTCCCATCACTTCCAAG 1766  
Db 1801 ATCAGAGAAATTAATTTGGAGCTTCTCAGACCGGATGACAAAAGTCCCATCTCTTCCAAG 1860  
QY 1767 AAACCGCATTTGACATAAGCAGAGTTTGGCATCACGGTTACAGATACCGAGATGGCTAC 1826  
Db 1861 AAACATGCTTTTGACATCACCAGAGCTTTCCATCATGTCTACAAATATTCAGATGCTTAC 1920  
QY 1827 AGCTTTGCCAAGTTGAAACAAAGAGTTTGGTGATGAGAACCGTCAATTGGAACCTGTGCT 1886  
Db 1921 ACTGTTTCCCAACACGAAACAAAGAAATTTGGTGATGATAAACCCTTCTTGAACCTCTCGCT 1980  
QY 1887 TTGTAACAACACTTCAAAATCTACAATATTAACTGAGG 1923  
Db 1981 TTGTAACAACATATAGAATGCATTAAATGTGGGAG 2017

RESULT 15  
AAF73373

ID AAF73373 standard; cdNA; 2089 BP.

XX AAF73373;

XX 30-APR-2001 (first entry)

XX Grand fir (-)-limonene synthase coding sequence SEQ ID NO: 5.

XX Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
KW terpinolene synthase; insect resistance; nutrition; ss.

XX Abies grandis.

XX WO200107565-A2.

XX 01-FEB-2001.

XX 24-JUL-2000; 2000WO-US020264.

XX 26-JUL-1999; 99US-00360545.

XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX Steele CL, Bohlmann J, Croteau RB, Phillips MA;

XX WPI; 2001-182782/18.

XX P-PSDB; AAB69372.

XX New nucleic acid encoding monoterpene synthases, for increasing terpene  
PT synthesis in plants, e.g. for increasing resistance to pests or for  
PT treatment of cancer.

Claim 33; Page 112-115; 175bp; English.

The present invention provides the protein and coding sequences of monoterpene synthases from the grand fir. These include (-)-camphene synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-limonene, (-)-alpha-pinene synthase, limonene synthase, myrcene synthase and pinene synthase. The sequences can be used to produce transgenic plants expressing high levels of the enzymes, resulting in levels which are useful in protecting against and treating cancers, and to confer insect resistance on plants

Sequence 2089 BP; 638 A; 415 C; 450 G; 586 T; 0 U; 0 Other;

Query Match 46.08; Score 925.4; DB 4; Length 2089;  
Best Local Similarity 70.68; Pred. No. 9.8e-252;  
Matches 1340; Conservative 0; Mismatches 506; Indels 51; Gaps 6;

Qy 72 TCGTGCCTCAGTTCTTCTCATGAGATTAAAGCTCTCCGTAGAACAAATCCCAACTCTTGGG 131  
Db 127 TCGTTGATCAGTTCCAGCAATGTSCAGAGGCTCTCTGTATCTCTACAGCAGTCCCAACA 186  
Qy 132 ATCTGAGGCGGGGAAATCGTTCGGCATTCCTAAACATGCTTTTGAACAAGCTCGCA 191  
Db 187 CTCAGATGCTAGGCGACAGAAAGCTCTGGTCAATCAACATGAATGACCACTGTATCC 246  
Qy 192 TCTACTCATTTCT-----GTACAGAGCGGTGGGCAACTATCATTCACAAC 236  
Db 247 CATCGTGATGATATGGTGGTGGTGTACTGCAAGACGCATAGCCGATCATTCACCAAC 306  
Qy 237 CTGTGGACGATGATTTTATACAGTCTCTGATCTCAACGGCTTATGAGCACCTGATTAC 296  
Db 307 CTGTGGGAAGATGATTTTATACAAATCATTTGCTCTCA---CCTTATGGGGATCTTCGTAC 363  
Qy 297 CGGGAAGCTCTGACAGACTTATTTGGGGAAGTAAAGGATAT-----AATGTTCAAT 347  
Db 364 AGTGAACGCTCTGAGACAGCTGTTTGGGAAGTAAAGAGATGTTTCAATCAATACCAAT 423  
Qy 348 TTCAAGTCGCTGGAAGATGGAGCAATGATCTCTTCAAGCACTTTTGGTGGTGCATGAC 407  
Db 424 AATAGAAATATTTGGTCCCAAAATGATCTCTTACAGCTTTGGATGGTGTATGAC 483  
Qy 408 GTTGAAGCTTTGGAAATTCGACAGCAATTTCAAAAAGAGATATAAAACGGCACTCGATTAT 467  
Db 484 AITGAAGCTCTGGGATAGATAGACATTTCCAAATGAGATAGAGTAGCCCTCGATTAT 543  
Qy 468 GTTAAACAGTTATTGGACGAAA---AAGGCAATTCGATGTGGGGAGAGTGTGTGACT 524  
Db 544 GTTTACAGTTATTGGAAGGAAAAGGAAGGCAATGGGTGGCAGAGATTTCTACTTTTCT 603  
Qy 525 GACCTCAACTCAACCGCTTTGGGCTTTCGAACCTCTCGACTTACACGGATACACTGTCT 584  
Db 604 GATCTCAACTCGACTGCTTTGGGCTTTCGAACCTCTCGACTGCAACGGATACAAATGTCT 663  
Qy 585 TCAGATGTTTGAACGTTTAAAGCAAAAATGGGCAATTTTCTCCACTGCGCAATATT 644  
Db 664 TCAGATGCTGGAATATCTTCAAGATGAAAAGGGGCAATTTGCTGCTGCTGCAATCCTA 723  
Qy 645 CAGATAGGGAGAGATTAGAGGCTTCTCAATTTTATTCAGGGCTCCCTGCTGCGCTTT 704  
Db 724 ACCGAGGACAGATCCTAGAGTGTCTAAATTTATATCGGGTTCCTGCTGCTGCTTT 783  
Qy 705 CCGGGCGAGAAAGTTATGATGAGCTGAAACATTTCTACAAATATTTAAAGAGAGCC 764  
Db 784 CCGGGCGAGAAAGTTATGGAAGAGGCTGAAATCTTCTCGGCATCTTTATTTGAAAAGTC 843  
Qy 765 CTGCAAAAGATTCCGGCATCCGATATCTTCTACTAGATACGGGAGCTTCTGGATAT 824  
Db 844 TTACAAAGATTCCGGTCTCCAAAT---CTTTACGAGAGATAGAAATATGTTTGGAAATAT 900  
Qy 825 GTTTGGCACACCAATTTGCCACGCTTGGAGCAAGGAAATACATGGACGCTTTGGACAG 884  
Db 901 GTTTGGCACACCAATTTGCCAGATTGGAGCAAGAAATATATTCAGGCTCTACAGCAG 960

Qy 885 CACACTAAAAATA-----AGAACGCCGCCGAGAAACATTTTGAAGACTT 926  
Db 961 AGCGGCTATGAAGCTTAAACAGAGATGCCATATATGAACATGAAGAGCTTTTACAACTT 1020  
Qy 927 GCAAAATGGAATTCATATATTTTCACTCTCTTCAAGAGAGAGAGTTTAAACATGTTTCC 986  
Db 1021 GCAAAATGAGATTCAATATCTTTCACTCTCTTTCACACTAAGAGAGAGTTTACATCTATCC 1080  
Qy 987 CGATGTTGGAAGACTCGGGTCTCTCGAGATGACCTTCTGTGACATCTGTGACATCTGTCAGTG 1046  
Db 1081 AGATGTTGGAAGAAATCAGGTTCTGCTCAACTGACITTTTACGGCATCTGTCAGTGAA 1140  
Qy 1047 TACTACGCTTTGGCTTCTCTGCTTTCGAGCTCAACATTTCTGATTTTCAAGCTCGC 1106  
Db 1141 TACTACATATGCACTTCTGCAITTTCTATGTTGCCAAAACATTTTCAAGATGAG 1200  
Qy 1107 TTTACCAAGATGCTCATCTTTTATCAGGTTCTTGTGACGACATGTACGACGTTCTCGGACA 1166  
Db 1201 TTTGTCAAAGTGTCTCATCTTTGTAAACAGTTCTCGATGATATATATGACACTTTTGAACA 1260  
Qy 1167 GTAGACGAGCTGGAACTCTTTTACAGCGACAAATTAAGAGATGGGATCGTCCGGATGGAA 1226  
Db 1261 ATGACCAACTCCAACTTTTACGATGCAATTAAGAGATGGGATTTGTCAACAGCAAG 1320  
Qy 1227 TGCCTTCCAGATATATGAAAGGAGTGTACATGATGTTTATCACACCGTAAATGAAATG 1286  
Db 1321 TGGCTTCCAGATATATGAAAGGAGTGTACATGAGCTTTGTATCAATGCAATTAATGAAATG 1380  
Qy 1287 GCTCGAGTGCAGAGAGGCTCAAGCCGAGACACGCTCAACTATGCAAGACAGGCTTGG 1346  
Db 1381 GTGAGAGAGCTGAGAGACTCAAGGCCGAGATATGCTCAACTATATTTCAAAATGCTTGG 1440  
Qy 1347 GAGCGGTGTTTGTATGCTATATGCAAGAAAGTGAAGTGGATCGCCACTGGTATCTGCC 1406  
Db 1441 GAAGCCCTATTTGATACCTTTTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1500  
Qy 1407 ACCTTTGAGAGTACTTTGGAGAAACGGGAAAGTTAGCTCTGCTATCCCATATGCGCACTG 1466  
Db 1501 ACCTTTGAGAGTACTTTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1560  
Qy 1467 CAACCAATTTGAGTGGACATCTTCTTCTGATCAGATCAGATCTTCAAGGAGTGAAGTTC 1526  
Db 1561 CAACCAATTTGAGTGGATGATCAGCTTCTTCTGATCAGATCAGATCAGATCAGATCAGAT 1620  
Qy 1527 CCATCGAAGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1586  
Db 1621 CCATCGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
Qy 1587 TACAAGCAGACAGGCGCTTGGAGAGAGAGTTCGTCTATATCATGTTTATATGAAAGAC 1646  
Db 1681 TACAAGCGGATAGGCGCTTGGAGAGAGAGTTCAGCTATATCGTGTATATGAAAGAC 1740  
Qy 1647 AATCCTGGATTAAGGAGAGAGTTCGTCTGATCATATCACTTCAATGATCAGGAGCGCA 1706  
Db 1741 CATCCTGGATTAAGGAGAGAGTTCGTCTGATCATATCAAGCCATGATCAGTGAAGTGA 1800  
Qy 1707 ATCAGAAATTAATTTGGGAGCTTCTAAAGCCAGCAACAGTGTTCCTCATCTTCAAG 1766  
Db 1801 ATCAGAAATTAATTTGGGAGCTTCTCAGACCGGATAGCAAAAGTCCCATCTTCTCAAG 1860  
Qy 1767 AACACGCAATTTGACATAGAGAGTTCGTCTGATCAGGTTTACAGATACGAGATGGTAC 1826  
Db 1861 AACATGCTTTTACATCAGCAGAGCTTTCATCATCTGTACAAATATCGAGATGGTAC 1920  
Qy 1827 AGCTTTGCCAAGCTTGAACAAAGAGTTCGTCTGATGAGAACCGCTCAATGAACTGTGCT 1886  
Db 1921 ACTGTTTCCAAACAAAGAAATTTGATGAAACCGCTTCTTGAACCTTCTGCT 1980  
Qy 1887 TTGTAAACAACTTCAATCTCAATATTAACAGG 1923  
Db 1981 TTGTAAACAACTATAGAAATGCAATTAATAATGTTGGAAG 2017

Search completed: July 8, 2004, 12:35:48  
Job time : 918.973 secs

**This Page Blank (uspto)**

CF477103 RTW3 5 A  
BQ196773 NXLV105\_B  
BX680641 BX680641  
BG526917 NXPV 057

Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACAGCTATGACC).

# FEATURES

source	1. .745	/organism="Pinus taeda"
		/mol_type="mRNA"
		/strain="CCLONES"
		/db_xref="taxon:3352"
		/clone="RTW3 5 A06 A022"
		/lab_host="DH10B-T1 phage-resistant E. coli"
		/clone_lib="Well-ordered loblolly pine roots W3"
		/notes="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Query Match	28.0%;	Score 564.2;	DB 14;	Length 745;
Best Local Similarity	84.8%;	Pred. No. 6e-90;		
Matches 632;	Conservative 0;	Mismatches 113;	Indels 0;	Gaps 0;

QY 1100 ACTCGGTTTACCAAGATGCTCATCTTATACAGGTTTTCGACGACATGTACGAGCTT 1159

DB 1 ACTCGGTTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60

QY 1160 CGGACAGTACAGAGTGAAGTCTTTCAGCGGACATTAAGAGTGGATCGTCCG 1219

DB 61 CGGAACACTCGAGGAGTGAAGTCTTTCAGCGGACATTAAGAGTGGATCGTCCG 120

QY 1220 GATGGAATGCTTCCAGAAATATGAAGAGGTGTACATGATGTTTATCACACCGTAAA 1279

DB 121 GACAGAGTGCCTCCAGAAATATGAAGAGTGTACATGATGTTTATCACACTGTAAA 180

QY 1280 TGAATGGCTGCGAGTGCAGAGAGGCTCAAGCGCGAGACACGCTCAACTATGCAAGACA 1339

DB 181 TGAATGCTCTCAGGAGGACAGAGGCTCAAGCGCGAGACACGCTCAACTATGTCGACA 240

QY 1340 GGCTTGGAGCGGTGTTTTCGATATGCGAGAGCAAGTGGATCGCTCGCTTGA 1399

DB 241 GGCTTGGAGGGAATATATTGATTCGTATATGCAAGAGCAAGTGGATCGCTCGCTGTA 300

QY 1400 TCTGCCACGTTTGGAGGAGTACTTGGAGAACGGGAAAGTGTAGCTCTGCTCATCGCCCATG 1459

DB 301 GGTGCCAACATTTGAGAGGAGTACTACGAGACGGGAAATAGCTCTGCTCATCGGTATC 360

QY 1460 CGCACTGCACCCCATTCCTGAGTGGACATCCCTTTCCTGATCAGATCCCTCAAGGAAGT 1519

DB 361 GGCATTGCAACCCCATTCCTGAGCGACATCCCTTTCCTGAGCAGCTCCTCAAGGAAGC 420

QY 1520 TGACTTCCCATCGAGCTCATGACTTGTATGTATCATCTCTCGATTAGAGGTGATAC 1579

DB 421 TGACATTCATCGAGGCTCAATGACTTGGCATCTGCCATCTTCGATTACGAGGAGATAC 480

QY 1580 ACGGTGCTTACAAAGCGACAGCGGCGCGTGGAGAGAGAGCTTCGTCTATATCATGTTATAT 1639

DB 481 CGCTGCTTACCAGCGGACAGCGGCGCGTGGAGAGAGAGCTTCGTCTATATCATGTTATAT 540

QY 1640 GAAAGACATCTCGATTAACGAGAGAGATGCTCTGAATCATATCACTTCATGATCAG 1699

DB 541 GAAAGACATCTCGAGCAACGAGAGAGATGCTCTCAATCATATCAACCGCATGATCAG 600

QY 1700 GGACGCAATCAGAGAAATTAATTTGGAGCTTCTTAAGCGACAGACAGTGTTCCTCATCAC 1759

DB 601 TGATGTAATTAAGAGTTAATTTGGAGCTTCTCAACCAACAGCAGCGTTCCTCATATC 660

QY 1760 TTCCAAAGAACACGCAATTTACATAAGCAGAGATTTGGCATCAGGTTACAGATACCGAGA 1819

DB 661 TGCCAAAGAACACGCTTTTGGCATTAGCAGAGCTTTCCATTATGGCTACAAATATCGAGA 720

QY 1820 TGGCTACAGCTTTGCCCAAGTTGAA 1844

DB 721 TGGCTACAGCTTTGCCAGCATTGAA 745

RESULT 2

BQ196773 697 bp mRNA linear EST 07-MAY-2003

LOCUS NXLV105\_B02 F NXLV (Nsf Xylem Late wood Vertical) Pinus taeda cDNA

DEFINITION clone NXLV105\_B02 5' similar to Arabidopsis thaliana sequence

At4g16730 limonene cyclase like protein see http://mips.gsf.de/proj/thal/db/index.html, mRNA sequence.

BQ196773.1 GI:20379276

ACCESSION BQ196773

VERSION Pinus taeda (loblolly pine)

KEYWORDS EST.

SOURCE Pinus taeda

ORGANISM Pinus taeda

REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. Sederoff, R.

AUTHORS Molecular Basis of Wood Formation in the Pine Megagenome

TITLE Unpublished (2000)

JOURNAL Contact: Sederoff, Ron

COMMENT Forest Biotechnology

North Carolina State University

840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh, NC 27695, USA

Tel: 919 515 7800

Fax: 919 515 7801

Email: ron.sederoff@ncsu.edu, jerri.johnson@ncsu.edu

Please see http://web.ahc.umn.edu/biodata/nsfpine/ for further information.

Seq primer: T3.

Location/Qualifiers

1. .697

/organism="Pinus taeda"

/mol\_type="mRNA"

/strain="Coastal plain loblolly pine from North Carolina"

/db\_xref="taxon:3352"

/clone="NXLV105\_B02"

/tissue\_type="primary xylem"

/dev\_stage="late wood"

/lab\_host="Xli-Blue"

/clone\_lib="NXLV (Nsf Xylem Late wood Vertical)"

/notes="Vector: pTriplex; Site 1: EcoRI; The library is from late (summer-August) wood, taken from below the crown of a 20 year old tree. The harvested xylem tissue was on the cusp between transitional and mature wood. NOTE: The sequences contain a 'cDNA adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'AATTCGGCAATTATGCCC'."

ORIGIN

Query Match 23.6%; Score 475.2; DB 13; Length 697;

Best Local Similarity 83.0%; Pred. No. 3.3e-74;

Matches 534; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 1250 AGTGTACATCATGTTTATCACACCGTAAATGCTCGAGTGCAGAGAGGCTCA 1309

DB 22 AGTTTACATCATGTTTATCAACACTGTAAATGCTCTCAGAGGACAGAGGCTCA 81

QY 1310 AGCCGAGACACGCTCAACTATGCAACAGAGGCTTGGAGCGTGTTCGTTTCGTATAT 1369

DB 82 AGCCGAGACACGCTCAACTATGCTCAGAGGCTTGGAGGAATATATTCGTTATAT 141

QY 1370 GCAGGAAGCAAGTGGATCGCCACTGTTATCTGCCCAACGTTTTCGAGGAGTCTTGGAGAA 1429

DB 142 GCAGGAAGCAAGTGGATCGCCCGTGGTGGAGTCCCAACATTTTCGAGGAGTACTATGAGAA 201

QY 1430 CGGGAAGTTAGCTCTGCTCATCGCCATGCGCATGCAACCCATCTGACCTTGACAT 1489  
 Db 202 CGGGAAGTTAGCTCTGCTCATCGCCATGCGCATGCAACCCATCTGACCTTGACAT 261  
 QY 1490 CCCCTTCTGATCATCTCAAGGAGTTGACCTCCATCGAAGCTCAATGACTTGTAT 1549  
 Db 262 CCCCTTCTGATCATCTCAAGGAGTTGACCTCCATCGAAGCTCAATGACTTGTAT 321  
 QY 1550 AFTGATCATCTTCCGATTAAGAGTGATACACGGTGCTACAGGCGAGACAGGCGCGTGG 1609  
 Db 322 ATCTGCCATCTTCTGATTAAGAGTGATACACGGTGCTACAGGCGAGACAGGCGCGTGG 381  
 QY 1610 AGAAGAAGCTTCTGATTAAGAGTGATACACGGTGCTACAGGCGAGACAGGCGCGTGG 1669  
 Db 382 AGAAGAAGCTTCTGATTAAGAGTGATACACGGTGCTACAGGCGAGACAGGCGCGTGG 441  
 QY 1670 TGCTCTGATCATCTCAAGGAGTTGACCTCCATCGAAGCTCAATGACTTGTAT 1729  
 Db 442 TGCTCTGATCATCTCAAGGAGTTGACCTCCATCGAAGCTCAATGACTTGTAT 501  
 QY 1730 TCTAAAGCCAGACAAAGTGTTCCCATCTTCCAGGAGACAGGCGAGACAGGCGCGTGG 1789  
 Db 502 TCTAAAGCCAGACAAAGTGTTCCCATCTTCCAGGAGACAGGCGAGACAGGCGCGTGG 561  
 QY 1790 AGTTTGGCATCAGGTTACAGTACCGAGATGGCTACAGCTTTGCGCAACAGTTGAAACAAA 1849  
 Db 562 AGTTTGGCATCAGGTTACAGTACCGAGATGGCTACAGCTTTGCGCAACAGTTGAAACAAA 621  
 QY 1850 GAGTTTGGTGATGAGAACCGTCAATGAACTCTGCTTTGTA 1892  
 Db 622 NNGTTTGGTGANNNACCGTCAATGATCTGCTCACTTATNA 664

RESULT 3  
 LOCUS BX680641 599 bp mRNA linear EST 28-OCT-2003  
 DEFINITION BX680641 RS Pinus pinaster cDNA clone RS46D06, mRNA sequence.  
 ACCESSION BX680641  
 VERSION BX680641.1 GI:38015099  
 KEYWORDS EST.  
 SOURCE Pinus pinaster  
 ORGANISM Pinus pinaster  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
 1 (bases 1 to 599)  
 Frigerio,J. and Plomion,C.  
 Identification of water-deficit responsive genes in Maritime pine  
 (Pinus pinaster Ait.) using an EST approach  
 Unpublished (2002)  
 CONTACT: Frigerio JM  
 Genetique et Amelioration 69  
 INRA  
 route d'Arcachon 33612 Cestas CEDEX France  
 Email: Frigerio@pierrot.inra.fr  
 Email: Frigerio@pierrot.inra.fr  
 Seq primer: f3.

FEATURES  
 source location/Qualifiers  
 1..599  
 /organism="Pinus pinaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:71647"  
 /clone="RS46D06"  
 /tissue\_type="root"  
 /dev\_stage="6 weeks old seedling"  
 /lab\_host="SOLR"  
 /clone\_lib="RS"  
 /note="Vector: Uni-ZAP XR; ecotype: Landes; The library was made from the roots of 6 weeks old seedlings grown in hydroponic conditions. A three weeks drought stress treatment was applied by lowering the osmotic potential of the nutrient solution to -0.45 MPa using 170 g/l of polyethylene glycol as an osmoticum. A mixture of

ORIGIN  
 Query Match 22.2%; Score 447.2; DB 13; Length 599;  
 Best Local Similarity 85.2%; Pred. No. 3.1e-69;  
 Matches 511; Conservative 0; Mismatches 88; Indels 1; Gaps 1;  
 QY 1228 GCCTTCCAGATATATGAAGAGTGATCATGATGTTTATCACACCGTAAATGAATGG 1287  
 Db 1 GCCTTCCAGATATATGAAGAGT-TACATGATGTTTACACACCGTAAATGAATG 59  
 QY 1288 CTCGAGTGGCAGAGAGGCTCAAGGCCGAGACACGCTCAACTATATCAAGACAGGCTTGG 1347  
 Db 60 CTGCGGAGGCAAGAGGCTCAAGGCCGAGACACCTCTCAACTATGCTCGACAGGCTTGG 119  
 QY 1348 AGGCGTGTGTTGATTCGTATATGAGGAGACGAAAGTGGATGCGCACTGGTTATCTGCCA 1407  
 Db 120 AGGATATCTTGTATTCGTATATGCAAGAGCAAGAGTGGATAGCCACGGGTTATCTGCCAA 179  
 QY 1408 CGTTTGAGGAGTACTTGGAGAACGGGAAAGTTAGCTCTGCTCATCGCCCATCGCACCTGC 1467  
 Db 180 GCTTCAGAGATATCTTGGAGAACGGGAAAGTTAGCTCTGCGGCATCGCGTCTCGGCGTTGC 239  
 QY 1468 AACCCATCTGACGTTGGACATCCCTTTTCCTGTATCATATCCTCAAGGAAGTTGACTTCC 1527  
 Db 240 AACCCATCTGACGATGACATCCCTTTTCCTGTATCATATCCTCAAGGAAGTTGACTTCC 299  
 QY 1528 CATCGAAGCTCAATGACTTGTATGATGATCATCTTCGATTAAGAGGTGATCACGGTGT 1587  
 Db 300 CATCCAACTCAATGACTTGGCATGTGCCATCTTCGATTACGAGGTGATACCCGATGTT 359  
 QY 1588 ACAAGGACAGACAGGCGCTGGAGAGAGGCTCTGCTATATCATGTTATATGAAGACA 1647  
 Db 360 ATCAGGAGACAGGCTCTGGAGAGAGGAACTCTGTTATATCTTGTATATGAAGACA 419  
 QY 1648 ATCTGTGATTAACGGAAGAGATGCTCTGAATCATATCAATCTCATGATCAGGGACGCAA 1707  
 Db 420 ACCCTGGAGCAACAGAGAGATGCTCTTAATCATCTCAATCTCATGATCAGTGGCGTAA 479  
 QY 1708 TCAGAGATTAATTTGGAGGCTTTAAAGCCAGACAAACAGTGTTCCTCCATCATCTTCCAGA 1767  
 Db 480 TTAAGAGATTAATTTGGAGGCTTTCAAACCCGACAAACAGTGTGCCCCATTTCTTCCAAA 539  
 QY 1768 AACACGCAATTTGACATAAGCAGAGTTTGGCATCAGGTTACAGATPACCGAGATGCTTACA 1827  
 Db 540 AAATTACTTTTGACATTACAGAGCTTTCCATTACGTTTACAATACCGAGATGCTTACA 599  
 RESULT 4  
 LOCUS BG526917 599 bp mRNA linear EST 07-MAY-2003  
 DEFINITION NXPV\_057\_D04\_F NXPV (Nsf Xylem Planings wood Vertical) Pinus taeda cDNA clone NXPV\_057\_D04\_5, similar to Arabidopsis thaliana sequence At4g16730 limonene cyclase like protein see http://mips.gsf.de/proj/thal/db/index.html, mRNA sequence.  
 ACCESSION BG526917  
 VERSION BG526917.1 GI:13536796  
 KEYWORDS EST.  
 SOURCE Pinus taeda (loblolly pine)  
 ORGANISM Pinus taeda  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
 1 (bases 1 to 599)  
 Sederoff,R.  
 Molecular Basis of Wood Formation in the Pine Megagenome  
 Unpublished (2000)  
 CONTACT: Sederoff, Ron  
 Forest Biotechnology  
 North Carolina State University  
 840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,  
 NC 27695, USA

genotypes were used. Oligo-dT primed cDNA was directionally cloned into the EcoRI-XhoI lambda-ZAP vector arms and mass-excised to form a pBluescript phagemid"

Tel: 919 515 7800  
Fax: 919 515 7801  
Email: ron.soderoff@ncsu.edu, jerri.johnson@ncsu.edu  
Please see <http://web.ahc.umn.edu/biodata/nsfpine/> for further information.  
Seq primer: T3

RESULT 5	
CF474786	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Research	1980, Vol. 83, No. 2, pp. 11-20
3. The Importance of Parental Involvement	Journal of Educational Research	1980, Vol. 83, No. 3, pp. 21-30
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 31-40
5. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 5, pp. 41-50
6. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 6, pp. 51-60
7. The Importance of Teacher Evaluation	Journal of Educational Research	1980, Vol. 83, No. 7, pp. 61-70
8. The Effect of Teacher Salary on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 8, pp. 71-80
9. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 9, pp. 81-90
10. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 10, pp. 91-100

An EST database from well-watered loblolly pine (*Pinus taeda*) roots. Unpublished (2002).  
Other\_ESTs: RWXW2\_7.B11.b1.A021  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia, plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACAGCTATGACC).

FEATURES  
source

## ORIGIN

	Query Match	21.2%;	Score 427;	DB 14;	Length 700;
	Best Local Similarity	75.7%;	Prod. No. 1.1e-65;		
	Matches 529;	Conservative	0; Mismatches 170;	Indels 0;	Gaps 0;
1034	TCGTCACGTGGGAATACTACGCTTTGGCTTCTCTGCATTTCGAGCTTCAGAGCCTCAACATTCCTGG	1093			
1	TCGTCACGTGGNAATCTACACTTTAGCAGCTTGCATTGCAATGATCTTAAACATTCCTGC	60			
1094	ATTGAGATTCGCGCTTTACAGATGCTCTCATCTTTATCAGGTTCTTTGACGACATGTACGA	1153			
61	GTTCGACTAGGANTTGGTAAATAAGTCATATGATCACGATTCGACGATATCTACGA	120			
1154	CGCTTTCCGCACAGTAGCAGGCTGGAACTCTTCACAGCGACAATTAAAGATGGGATCC	1213			
121	CACCTTCGGAAACAATGGAGGAGCTCGAACTCTTAAACCGACGCTTTAAAGAGATGGGATCC	180			
1214	GTCCGGATGGNAATGCCCTTCGAAATATATGAAGGAGGTGTACATGATGGTTTATCACAC	1273			
181	GTCTTCGATAGAGTGTCTTCAGATTATATGAAGGAGGTGTACATGGCGGGTTACGACAA	240			

1274	QY	CGTAAATGAAATGGCTCGATGGCAGAGAAGCGCTCAAGGCGGAGACACGCTCAACTATGC	1333
241	Db		
241	Db	CATCAACGAAATGGCAGCAGAGGCGCAGAAAAATCAAGGGTGGGATACAGTCAGCTATGC	300
1334	QY	AAGACAGCGTTGGGAGGCGTGTTTTGATTCGTATATCGAGAGACCAAGTGGATCGGCAC	1393
301	Db	TCGAAATCTTGGAGGCTTTATTTGGTGTCTTATACAGAAGCCCAAGTGGATTTCAG	360
1394	QY	TGTTATCTGCCACGTTTGAGGAGTAGTACTTGGAGAAGCGGAAAAGTTAGCTCTGCTCATCG	1453
361	Db	TGTTATCTTCCCAAGTTTCGACGAGTAGTACCTCGAGAATTGGGAAGGTCAGCTTCGGCTCTCG	420
1454	QY	CCCATGGCACTGCAACCCATTCTGACGTGGACATCCCTCTTCTGATCACATCTCTCAA	1513
421	Db	CATAAACACCGCTCGAACCCATGCTGCACTTTGGGGTTTTCTCTTCCGCGCTCGAATCTCTGCA	480
1514	QY	GGAAATTGACTTCCCATCGAAGCTCAATGACTTCATATGTATCATCTCTCGATTAAAGAGG	1573
481	Db	GGAAATTGACTTTCCTCATCGAAATTCAATGATTTGATATGTGCCATCTTTCGACTGAAGG	540
1574	QY	TGATACACGGTGTCTAACGGCAGACACGGCCCGTGGAGAGAGAGCTTCGTCTATATCATG	1633
541	Db	TGACACTCAATGTCTACAAGGCTGACAGGGCGCGTGGAGAAGAGCTTCGCGCGGTATCGTG	600
1634	QY	TTATATGAAAGACAATCCTGGATTAAACGGAAGAGATGCTCTGAATCATATCAACTTTCAT	1693
601	Db	TTATATGAAAGACCATCTCTGGAATAACAGAGGAAGATGCTGTCTCATCAAGTCAATGCTAT	660
1694	QY	GATCAGGGAGCGAATCAGAGAAATTTAAATTTGGAGGCTTCT	1732
661	Db	GGTCGATTAACCTTACCAAGGAACCTGAATGGGAGTTACT	699

RESULT 6	CF476978	LOCUS	CF476978	637 bp	mRNA	linear	EST 08-SEP-2003
DEFINITION	RTW3_5_A06_b1_A022 Well-watered loblolly pine roots WW3 Pinus						
ACCESSION	taeda cDNA clone RTW3_5_A06_A022 3', mRNA sequence.						
VERSION	CF476978						
KEYWORDS	CF476978.1 GI:34505847						
	EST.						

SOURCE	Pinus taeda (loblolly pine)
ORGANISM	Pinus taeda
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 637)
AUTHORS	Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and Neale, D.
TITLE	An EST database from well-watered loblolly pine ( <i>Pinus taeda</i> ) roots
JOURNAL	Unpublished (2003)
COMMENT	Other ESTs: RTW3 5 A06.G1 A022 Contact: Cordonnier-Pratt MM Laboratory for Genomics and Bioinformatics The University of Georgia, Department of Plant Biology Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA Tel: 706 542 1860 Fax: 706 583 0210

Email: mmpratt@uga.edu  
RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below 5' end quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: M13-21 (TGTAAGACGACGCCAGT)  
POLYA=Yes.

```
FEATURES
  source
    Location/Qualifiers
      1..637
        /organism="Pinus taeda"
        /mol type="mRNA"
```

```

/strain="CCLONES"
/db xref=taxon:3352"
/clone="RTW3_5_A06_A022"
/lab_host="DH10B-T1 phage"
/clone_lib="Well-watered"
/note="vector: pSL1180; S
library was prepared from
(Pinus taeda) roots water
day. Pre-dawn water poten
Roots were harvested for
cDNA was cloned unidirect
excised with EcoRI (5', en

```

## ORIGIN

Query Match	21.2%	Score 425.8	DB 14	Length 637
Best Local Similarity	85.3%	Pred. No. 1.8e-65		
Matches 475	Conservative 0	Mismatches 82	Indels 0	Gaps 0
QY	1336	GACAGGCTTGGGAGGCGTGTGTTTGAATTCGTTATATGACGAAAGCAAAAGTCGATCGCACATG	1395	
Db	1	GACAGGCTTGGGAGGAAATTAATTTGATTCGTTATATGCAAGAAGCAAAAGTGGATCGCCAGTG	60	
QY	1396	GTTATCTGCCACGTTTGAGGAGTACTTTGGAGAACGGGAAAGTTAGTCTCTGTCATCGCC	1455	
Db	61	GTGAGGTGCCAACATTTGAGGAGTACTACGAGAAACGGGAAAAATTAGCTCTGGTCATCGCG	120	
QY	1456	CATGCCACTGCNAACCCATTCTGACGTTGGACATCCCTTCTCTGATCATCATCTCAAGG	1515	
Db	121	TATCGGCATTTGCAACCCATTCTGACACCGACATCCCTTCTCTGACACGTCCTCAAGG	180	
QY	1516	AAAGTTGACCTTCCCATCGAAGCTCAATGACTTGATATGTATCATCTTCGATTAAAGAGTG	1575	
Db	181	AAAGTCGACATTCATCGAAGCTCAATGACTTGGCATCTGCCATCTTCGATTACGAGGAG	240	
QY	1576	ATACAGGTGCTTACAGGACAGACAGGCCCGTGGAGAAAGGTTCTGCTATATCATGTT	1635	
Db	241	ATACGCGCTGCTACCAAGGCGGACAGGCGCCGTGGAGAAAGGCTTCGTGTATATCTGTT	300	
QY	1636	ATATGAAGACAAATCTCTGATTAACGGAAGAGATGCTCTGAATCATATCAACTTCATGA	1695	
Db	301	ATATGAAGACAAATCTCTGAGCAACGGAAGAGATGCTCTCAATCATATCAACGCCATGA	360	
QY	1696	TCAGGACGCAATTCAGAGAAATTAATTTGGAGGCTTCTAAAGCCAGACAAACAGTGTCCCA	1755	
Db	361	TCAGTGATCTAATTTAAAGGATTAATTTGGAGCTTCTCAACCAACACAGCGGTTCCCA	420	
QY	1756	TCACITCCAAGAAACACGCGATTGACATAAGCAGAGTTTGGGATCATCGGTTACAGATACC	1815	
Db	421	TATCTGCCAAAAACAATGCTTTTGACATTTAGCAGAGCTTTCCTATTTATGCTACAAATATC	480	
QY	1816	GAGATGGCTTACAGCTTTTGCCACGTTTGAAACAAAGAGTTTGGTGATGAGAACCGTCATTG	1875	
Db	481	GAGATGGCTTACAGCGTTTGCCAGCATTTGAACAAAGAGTTTGGTGAACGCAACCGTCATTG	540	
QY	1876	AACCTGTGCTTTGTAA	1892	
Db	541	ATGCTGTGACTTTATAA	557	

RESULT 7	CF401916	LOCUS	CF401916	696 bp	mRNA	linear	EST 29-AUG-2003
DEFINITION	RTW1_15_B05_g1_A015 well-watered loblolly pine roots WW1 Pinus taeda cDNA clone RTW1_15_B05_A015 5', mRNA sequence.	ACCESSION	CF401916				
VERSION	CF401916.1	KEYWORDS	GI:34360333				
SOURCE	Pinus taeda (loblolly pine)	ORGANISM	Pinus taeda				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 696)	AUTHORS	Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J.				

Gebredmedhin, M., Dervinis, C., Martin, T., White, T., Davis, J., and Neale, D.  
An EST database from well-watered loblolly pine (Pinus taeda) roots  
Unpublished (2003)  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmp@prattuga.edu  
RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACAGCTATGACC).

FEATURES

source  
1..696  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="CCLONES"  
/db\_xref="taxon:3352"  
/clone="RTW1 15 B05 A015"  
/lab\_host="DH10B-T1 Phage-resistant E. coli"  
/clone\_lib="Well-watered loblolly pine roots WW1"  
/notes="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/- 0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Query Match 20.8%; Score 419.2; DB 14; Length 696;  
Best Local Similarity 75.1%; Pred. No. 2.6e-64;  
Matches 523; Conservative 0; Mismatches 173; Indels 0; Gaps 0;  
1065 TGCATTGCGTTGAGGCTCAACATCTGGATTCAGACTCGGTTTACAGATGTCTCAT 1124  
Db 1 TGCATTGCAATGATCTCAACATCTTGGCTTCGACTAGGATTTGTAATAAGTCA 60  
1125 CTATCAGGTTCTTGACGACATGACAGCTCTTCGGCAGTAGAGCTGGACTC 1184  
Db 61 ATGATCAGGATCTCGACGATATCTAGCACCTTCGGAACAAATGGAGGAGCTCGAATC 120  
1185 TTCACAGCGACAATTAGAGATGGATCCGTCGCGATGGAATGCCCTTCCAGAAATATG 1244  
Db 121 TTAACCCGAGGTTTAAGAGATGGATCCGTCCTTCGATAGAGTGTCTCCAGATTATG 180  
1245 AAGGAGTGTACATGATGTTTATFACACCGTAAATGAATGGCTCGAGTGGCAGAGAAG 1304  
Db 181 AAGGAGTGTACATGATGCGGTTTACGACACATCAACGAAATGGCAGAGGCGCAGAAA 240  
1305 GCTCAAGCCGAGACAGCTCAACTATGCAAGACAGAGCTGGGAGGCGTGTGTTGATCG 1364  
Db 241 ATTCAAGCTGGGATACAGTCAGTATGCTCGAATAATCTTGGAGGCTTTATTTGGTCT 300  
1365 TATATCAGGAAGCAAGTGGATCGGCATCGGTTATCTGCCACGTTTGGAGGAGTACTTG 1424  
Db 301 TATATCAAGGAAGCAAGTGGATTTCCAGTGGTTATCTTCCACGTTTCGACGAGTACCTC 360  
1425 GAGAACGGAAAGTAGCTGTGTCATCGCCCATCGCAGTCAACCCATCTTCACGTTG 1484  
Db 361 GAGAACGGAAAGTGTGCTGCTCGGCTCTCGATACACCAACCGTTCGAACCCATGCTGACTTTG 420  
1485 GACATCCCTTCTCTGTATCATCTCTCAAGGAAGTGTGACTTCCCATCGAAGCTCAATGAC 1544  
Db 421 GGGTTTCTCTTCGCTCGATCTCTCGAGGAATTCATTTCCATCGAATTCATGAT 480

1545 TTGATATGATATCATCTCTCGATTAAAGAGGTGATACAGGTGCTGCTCAAGGCAGACAGGCC 1604  
Db 481 TTGATATGTCCTATCTCTCGACTGAAAGGTGACACTCAATGCTTACAGGCTGACAGGCG 540  
1605 CGTGGAGAGAGAGCTTCGCTCTATATCATGTTATATGAAGACAATCTCTGATTAACGGAA 1664  
Db 541 CGTGGAGAGAGAGCTTCGCGCGTATCGTGTATATGAAGAGACCATCTCTGGAATTAACAGAG 600  
1665 GAAGATGCTCTGAATCATATCACTTCATGTCAGGAGCGCAATCAGAGAAATTAATTTGG 1724  
Db 601 GAAGATGCTGTCTCAATCAAGTCAATGCTATGCTGATTAACCAAGGACTGAATTTGG 660  
1725 GAGCTTTCTAAAGCCAGACACAGTGTCCCATCACT 1760  
Db 661 GAGTACTTAGACCCGACAGCGGTGTTCCTCATCTCT 696

RESULT 8  
CF479802/c  
LOCUS  
DEFINITION  
CF479802 669 bp mRNA linear EST 08-SEP-2003  
taeda cDNA clone RTW3\_12\_D02\_A022 3', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Pinus taeda (loblolly pine)  
Pinus taeda  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

REFERENCE  
AUTHORS  
Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.M., Dean, J.,  
Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J.,  
Cannon, R., Owen, A. and Neale, D.  
EST database from well-watered loblolly pine (Pinus taeda) roots  
Unpublished (2003)

TITLE  
JOURNAL  
COMMENT  
Other ESTs: RTW3\_12\_D02.g1\_A022  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmp@prattuga.edu  
RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: M13-21 (TGTAACGACGCGCAGT)  
POLYA=No.

FEATURES

source  
1..669  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="CCLONES"  
/db\_xref="taxon:3352"  
/clone="RTW3\_12\_D02\_A022"  
/lab\_host="DH10B-T1 Phage-resistant E. coli"  
/clone\_lib="Well-watered loblolly pine roots WW3"  
/notes="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/- 0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Query Match 20.5%; Score 413; DB 14; Length 669;  
Best Local Similarity 76.1%; Pred. No. 3.3e-63;  
Matches 509; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

QY 994 GGAAGACTCGGGTCTCTGAGATGACCTTCTGTGCGACATCGTCACGTGGGAATCTACG 1053  
Db 669 GGAAGATATGATGATTCCTGAAATAAATCTCAGCGCATCGTCACGTGGGAATCTACA 610  
QY 1054 CTTTCGGTCTCCGATTCGGTTCGAGCCTCAACATCTGGATTTCAGATCGGCTTTACCA 1113  
Db 609 CTTTACGAGCTTGCATTCGAAATGATCCCTAAACATCTCGGTTTCGATAGGATTTGGTA 550  
QY 1114 AGATGCTCATCTATCATCGGTTCTTGGACGACATGTACGACCTCTTCGGSCACAGTAGAG 1173  
Db 549 AAATAAGTCATATGATCAGATTCCTCGACGATATCTACGACACCTTCGGAACAATGGAG 490  
QY 1174 AGCTGGAATCTTTCACAGGCAATTAAGAGATGGATTCGTCGCGATGGGAATGCCCTTC 1233  
Db 489 AGCTCAAACTCTTAACCGCAGCGTTTAAAGATGGATTCGCTCTTCGATAGATGCTCTTC 430  
QY 1234 CAGAAATATGAAGAGGTATACATGATGGTTTATACACCGTAAATGAATGGCTCGAG 1293  
Db 429 CAGATTATGAAGAGGTATACATGGCGGTTTACGACAAATCAACGAAATGGCAGAG 370  
QY 1294 TGGCAGAGAAGCTCAAGCCGAGACAGCTCAACTATGCAAGACAGGCTTGGGAGGGT 1353  
Db 369 AGGCGCAGAAATTCAGGCTGGGATACAGTCAGCTATGCTCGAAATCTTGGGAGGCT 310  
QY 1354 GTTTTGATTCGATATGAGAGGAAAGTGAATGATCGCACTGTGTTATCTGCCACGTTG 1413  
Db 309 TTATTTGGTGCTTATATACAAGAAAGCAAGTGAATTCAGATGTTATCTCCACGTTG 250  
QY 1414 AGGAGTACTTGAGAACGGGAAAGTTAGCTGTCTATCGCCATGCGCACTGCAACCCA 1473  
Db 249 ACGAGTACTCGAATGGAAGGTGAGCTTCGGCTTCGGATACCACTGCAAGCTCGAACCA 190  
QY 1474 TTCTGACGTTGACATCCCTTCTGATCATCATCTCAAGAAAGTTGACTTCCCATCGA 1533  
Db 189 TGCTGACTTTGGGTTCTCTCTCGCCTCGAATCTCGCAGGAATTCGACTTTCATCGA 130  
QY 1534 AGCTCAATGACTTGATGATATCATCTTCGTTAAAGAGGTGATACAGGTGCTACAGG 1593  
Db 129 AATTCATGATTTGATATGTGCTATCTTCGACTGAAAGGTGACACTCAATGCTACAAG 70  
QY 1594 CAGACAGGCGCGTGGAGAAAGGCTTCGTCTATATCATGTTATGAAGACATCTG 1653  
Db 69 CTGACAGGCGCGTGGAGAAAGGCTTCGGCGTATCGTTGTTATGAAGACCATCTG 10  
QY 1654 GATTAACGG 1662  
Db 9 GAATAACAG 1

RESULT 9  
CF663845  
LOCUS  
DEFINITION  
RTCNT1\_5\_B08\_g1\_A029 Root control Pinus taeda linear EST 07-OCT-2003  
RTCNT1\_5\_B08\_A029 5', mRNA sequence.  
ACCESSION  
CF663845  
VERSION  
CF663845.1 GI:37561088  
KEYWORDS  
EST.  
SOURCE  
Pinus taeda (loblolly pine)  
ORGANISM  
Pinus taeda  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
REFERENCE  
1 (bases 1 to 616)  
Pratt,J.L., Cordonnier-Pratt,M.-M., Lorenz,W.W., Zimmermann,C. and  
Dean,J.F.D.  
An EST database from untreated loblolly pine (Pinus taeda) roots  
Unpublished (2003)  
Other ESTs: RTCNT1\_5\_B08.bl\_A029  
CONTACT: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860

Fax: 706 583 0210  
Email: mmpratt@uga.edu  
RNA prepared and library constructed by W. Walter Lorenz (School of  
Forest Resources, University of Georgia); plant material prepared  
by Craig Zimmermann (School of Forest Resources, University of  
Georgia) using rooted cuttings provided by the Forest Biology  
Research Cooperative (FBRC) and the CCLONES project a the  
University of Florida; sequencing done in the Laboratory for  
Genomics and Bioinformatics, University of Georgia. Sequence ends  
have been trimmed to exclude vector and regions below Phred quality  
16. Three-prime sequences are presented as their reverse complement  
and have been trimmed to exclude polyA.

Seq primer: JENREV (CAGGAAACAGCTATGACC).

#### FEATURES

source  
1..616  
Location/Qualifiers  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strains="3 CCLONES"  
/db\_xref="taxon:3352"  
/clone="RTCNT1\_5\_B08\_A029"  
/lab\_host="DH10B-T1 Phage-resistant E. coli"  
/clone\_lib="Root control"  
/note="Organ: root; Vector: pSL180; Site\_1: EcoRI;  
Site\_2: XhoI; The library was prepared from polyA+ RNA  
from the roots of 1-year-old loblolly pine (Pinus taeda).  
cuttings that were rooted and then planted in washed sand.  
Just before harvesting roots for RNA isolation, the rooted  
cuttings were maintained for 27 days (April 2003) under  
ambient conditions in a local greenhouse. They were kept  
on a weekly regimen of 0.5x nutrient-complete Hoagland's  
solution and supplemented with additional water sufficient  
to maintain a 15% soil moisture content. Double-stranded  
cDNA was cloned unidirectionally into pSL180. Inserts can  
be excised with EcoRI (5' end) and XhoI (3' end)."

#### ORIGIN

Query Match 19.9%; Score 399.6; DB 14; Length 616;  
Best Local Similarity 78.7%; Pred. No. 7.8e-61;  
Matches 477; Conservative 0; Mismatches 129; Indels 0; Gaps 0;  
QY 890 TAAATAAAGAACGCCGCCGAGAACTTTTAAAGACTTTCGAAATTCGAAATTCATATATT 949  
Db 11 TAAGACGACGACCAAGACCCGAGAACTTTTAAAGACTTTCGAAATTCGAAATTCATATCTT 70  
QY 950 TCACCTCCCTTACAGAGAGAGAGTTAAACACATGTTTCCCGATGCTGGAAGACTTCGGGTT 1009  
Db 71 TCACCTCCCTTACAGCAAAAAGAGTTAAACACGCTGTCAGATGGTGGAAAGATTCCGGTTT 130  
QY 1010 TCCTGAGATGACCTTCTGTGACATCGTTCAGTGAATATACGTTTGGTTCCTGCTCAT 1069  
Db 131 CTCTCAACTGACATTCATCTGGCATCGTCAGTGAATTCCTACACTTTGGCTCCTGCTCAT 190  
QY 1070 TGGGTTTCGAGCTCAACATCTTCGATTCAGACTCGGCTTTTACCAAGATGTCATCTTAT 1129  
Db 191 TGGCAGTGGCCCAACATTCAGCAATTCAGATGGGCTTTTCCCAAAACGTTTATCTTTG 250  
QY 1130 CACGGTTTCTTGACGACATGTACGACGCTTTCGGCACAGTAGACGAGCTGGAACCTCTTCA 1189  
Db 251 AATAGTTCTTGACGACATCTATGACATTCGGAACATGGAGGAGCTCGNACTCTTCAC 310  
QY 1190 AGCGACAAATTAAGAGATGGGATCCGTCGGGATGGAAATGCTTCGAAATATATGAAAGG 1249  
Db 311 AGCGCAATTAAGAGATGGGATCCGTCGGCAGGAGTTCCTTCCGAAATATATGAAAGG 370  
QY 1250 AGTGTACATGATGGTTTATCACACCGTAAATGAATGCTCGAGTGGCAGAGAGGCTCA 1309  
Db 371 CATATATATGATGGTTTATACGATTGCGTTAATCAAAATGGCTCGAGAGCGGAGAGAGTCA 430  
QY 1310 AGGCCGAGACACGCTCAACTATGCAAGACAGGCTTTGGGAGGCGTGTTCGATTCGTTAT 1369  
Db 431 AGGCCGAGACACGCTCACTACGCTCGAATACGTTGGGAGGCGGCTATTTGATGCTTCT 490  
QY 1370 GCAGGAAGCAAGTGGATCGGCATCTGCTTCTGCCACGTTTGGAGGAGTACTTTGGAGAA 1429

Db 491 GGAAGAACCAAGTGTCCAGCGTTATATACCAAGTTTGAGGATATTTCAGAA 550  
 QY 1430 CGGGAAGTAGCTCTGCTATCGCCATCGGACATCGCAACCAATTCAGCTTGACAT 1489  
 Db 551 CGGGAAGGTGAGTTTGGGTATCGAGCAGCAACATCGCAACCAATTCAGCTTGATAT 610  
 QY 1490 CCGCTT 1495  
 Db 611 TCCCT 616

RESULT 10  
 BQ698077  
 LOCUS  
 DEFINITION  
 NXPV\_064\_C05\_F\_NXPV (Nsf Xylem Planings wood Vertical) Pinus taeda  
 cDNA clone NXPV\_064\_C05\_5' similar to Arabidopsis thaliana sequence  
 At1g61680 hypothetical protein see  
 http://mips.gsf.de/proj/thal/db/index.html, mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Pinus taeda (loblolly pine)  
 Pinus taeda  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
 Sederoff, R.  
 1 (bases 1 to 516)

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Molecular Basis of Wood Formation in the Pine Megagenome  
 Unpublished (2000)  
 Contact: Sederoff, Ron  
 Forest Biotechnology  
 North Carolina State University  
 840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,  
 NC 27695, USA  
 Tel: 919 515 7800  
 Fax: 919 515 7801  
 Email: ron.sederoff@ncsu.edu, jerri.johnson@ncsu.edu  
 Please see http://web.ahc.umn.edu/biodata/nsfpine/ for further  
 information.  
 Seq primer: T3.

FEATURES  
 source  
 1..516  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="Coastal plain loblolly pine from North Carolina"  
 /db\_xref="taxon:3352"  
 /clone="NXPV\_064\_C05"  
 /tissue\_type="Xylem"  
 /cell\_type="Planings (secondary)"  
 /dev\_stage="Transitional"  
 /lab\_host="X11-Blue"  
 /clone\_lib="NXPV (Nsf Xylem Planings wood Vertical)"  
 /note="Vector: Bluescript SK; Site 1: Eco RI; Site 2:  
 XhoI; The library is from early (spring) secondary wood,  
 taken from a ten year old tree in the transitional phase.  
 The tree is a kind gift of the Westvaco Corporation.  
 Secondary xylem was harvested from the tree by peeling  
 back the bark and primary xylem and then removing the  
 underlying tissue with a block plane. NOTE: The sequences  
 contain a 'cDNA adapter' between the EcoRI site and the  
 start of the EST. The adapter sequence is  
 'AATTCGGCACGAG'."

ORIGIN  
 Query Match 19.3%; Score 388.6; DB 13; Length 516;  
 Best Local Similarity 84.7%; Pred. No. 7.3e-59;  
 Matches 436; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 1353 TGTATTGATTCGATATCGAGAGCAAGTGCATCGCCATCGTGTATCTGCCACGTTT 1412  
 Db 2 TATATTGATTCGATATCGAGAGCAAGTGCATCGCCATCGTGTAGGTGCCAACATTT 61

QY 1413 GAGGAGTACTTGGAGAACGGGAAAGTTAGCTCTGCTCATGCCCATCGGCACATGCAACCC 1472  
 Db 62 GAGGAGTACTATGAGAACGGGAAAGTTAGCTCTGCTCATGCCCATCGGCACATGCAACCC 121  
 QY 1473 ATTCTGACGTTGGACATCCCTTTCTGATCATCATCTCAAGGAAGTTGACTTCCCATCG 1532  
 Db 122 ATTTTGACGACCGACATCCCTTTCTGAGCAGCTCTCAAGGAAGTTGACATTCATCG 181  
 QY 1533 AAGCTCAATGACTTGTATATGTATCATCTTCGATTAAAGAGGTGATACACGGTCTACAAG 1592  
 Db 182 CAGTCAATGACTTGGCATCTGCCATCTTCGATTACGAGGGATACGGCTGCTACCAG 241  
 QY 1593 GCAGACAGGCGCGTGGAGAAAGCTTCGTCTATATCATGTTATATGAAGCAATCTCT 1652  
 Db 242 CGGACAGGCGCGTGGAGAAAGCTTCGTGTATATCTTGTATATGAAGCAATCTCT 301  
 QY 1653 GGATTAACGGAGAGATGCTCTGAATCATATCAACTTCATGATCAGGAGCGCAATCAGA 1712  
 Db 302 GGAACAACAGAGAGATGCTCTCAATCATCTCAAGCCCATGATCATGATGATTAATAAA 361  
 QY 1713 GAATTAATTTGGAGCTTCTAAAGCCAGACAACAGTGTTCCTCATCTTCCAAAGAACAC 1772  
 Db 362 GGATTAATTTGGAGCTTCTCAACCAACAGCAGCGTTCCTCATATCTGCCAAAACAT 421  
 QY 1773 GCATTTGACATAAGCAGAGTTTGGCATCAGCTTACAGATACCGAGATGCTACAGCTTT 1832  
 Db 422 GCTTTTGACATTAGCAGAGCTTTCCATTTGGTGTGCTACAAATATCGAGATGCTACAGGTT 481  
 QY 1833 GCCAAGCTTGAACAAGAGTTTGGTGATGAGAAC 1867  
 Db 482 GCCAAGCTTGAACAAGAGTTTGGTGAGAGAAC 516

RESULT 11  
 BQ677624  
 LOCUS  
 DEFINITION  
 BQ677624 RN Pinus pinaster cDNA clone RN42B08, mRNA sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 EST.  
 SOURCE  
 Pinus pinaster  
 Pinus pinaster

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 1 (bases 1 to 517)  
 Frigerio, J. and Plomion, C.  
 Identification of water-deficit responsive genes in Maritime pine  
 (Pinus pinaster Ait.) using an EST approach  
 Unpublished (2002)  
 Contact: Frigerio JM  
 Genetique et Amelioration 69  
 INRA  
 route d'Arcachon 33612 Cestas CEDEX France  
 Email: Frigerio@pierrot.inra.fr  
 Email: Frigerio@pierrot.inra.fr  
 Seq primer: T3.

FEATURES  
 source  
 1..517  
 /organism="Pinus pinaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:71647"  
 /clone="RN42B08"  
 /tissue\_type="root"  
 /dev\_stage="6 weeks old seedling"  
 /lab\_host="SOLR"  
 /clone\_lib="RN"  
 /note="Vector: Uni-ZAP XR; ecotype: Landes; The library  
 was made from the roots of 6 weeks old seedlings grown in  
 hydroponic conditions. A mixture of genotypes used.  
 Oligo-dT primed cDNA was directionally cloned into the  
 EcoRI-XhoI lambda-ZAP vector arms and mass-excised to form  
 a pBluescript phagemid"

ORIGIN



Db 658 AGAGGAGAGATGAGAGATGCTCAATTAATTCGGCATCTCTCATTTGCCCTTCTCG 717  
 QY 710 CGAGAAAGTATGATGAGAGCTGAACATCTCTCAAAAATATTAAAGAGAGCCCTGCA 769  
 Db 718 CGAAAAAATATGAGAGCGGCTGAATCTCTCTACAATGTATTTAAAAAGATGCCCTACA 777  
 QY 770 AAGATTCCGGCATCCAGTAT 790  
 Db 778 AAGATTCCGCCCTCCGGTCT 798

RESULT 13  
 LOCUS AL750951 481 bp mRNA linear EST 20-JUN-2002  
 DEFINITION AL750951 RS Pinus pinaster cDNA clone RS02D01 similar to PINENE SYNTHASE, mRNA sequence.  
 ACCESSION AL750951  
 VERSION AL750951.1 GI:21492198  
 KEYWORDS EST.  
 SOURCE Pinus pinaster  
 ORGANISM Pinus pinaster  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 481)  
 REFERENCE Frigerio, J. and Plomion, C.  
 AUTHORS Identification of water-deficit responsive genes in Maritime pine (Pinus pinaster Ait.) using an EST approach  
 TITLE Unpublished (2002)  
 JOURNAL Contact: Frigerio JM  
 COMMENT Genetique et Amelioration 69  
 INRA  
 route d'Arcachon 33612 Cestas CEDEX France  
 Email: Frigerio@pierrotton.inra.fr  
 Seq primer: T3.

FEATURES  
 source Location/Qualifiers  
 1..481  
 /organism="Pinus pinaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:71647"  
 /clone="RS02D01"  
 /tissue\_type="root"  
 /dev\_stage="6 weeks old seedling"  
 /lab\_host="SOLR"  
 /clone\_lib="RS"  
 /note="Vector: Uni-ZAP XR; ecotype: Landes; The library was made from the roots of 6 weeks old seedlings grown in hydroponic conditions. A three weeks drought stress treatment was applied by lowering the osmotic potential of the nutrient solution to -0.45 MPa using 170 g/l of polyethylene glycol as an osmoticum. A mixture of genotypes were used. Oligo-dT primed cDNA was directionally cloned into the EcoRI-XhoI lambda-ZAP vector arms and mass-excised to form a pBlue-script phagemid"

ORIGIN  
 Query Match 18.7%; Score 376; DB 9; Length 481;  
 Best Local Similarity 86.5%; Pred. No. 1.3e-56;  
 Matches 415; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
 QY 1038 CAGTGGAATACACGCTTTGGCTTCTGATTCGGTTCGAGCCTCAACATTCGGATTC 1097  
 Db 2 CAGTGGAATACACGCTTTGGCTTCTGATTCGGTTCGAGCCTCAACATTCGGATTC 61  
 QY 1098 AGACTCGGCTTACCAAGATGCTCATCTTATACGGTTCCTGACGACATGTACGAGTC 1157  
 Db 62 AGACTCGGCTTTCGGAAGCGTGTCATATATCACTGTTCTCGACGATGTACGACCTC 121  
 QY 1158 TTCGGCACAGTAGACGAGCTGGAATCTTTACACGCGACAATTAAGATGGGATCCGTC 1217  
 Db 122 TTCGGAACAGTTGATGAGCTCAAACTGTTACACCGCGCAATTAAGAGATGGATCCGTC 181  
 QY 1218 GCGATGGAATGCCCTCCAGAAATATATGAAGAGGTGTACATGATGGTTTATCACACCGTA 1277

Db 182 GCCACAGATTGCTTCCACAATATATGAAAGGAATTTACATGATGGTTTATCAACACCGTA 241  
 QY 1278 AATGAATGCTCGAGTGGCAGAGAGGCTCAAGCGCGAGACACGCTCAACTATGCAAGA 1337  
 Db 242 AATGAATGCTCGGAGGACAGAAAGGCTCAAGCGCGAGACACCTCAACTATGTCGA 301  
 QY 1338 CAGGCTTGGGAGCGTGTGTTTATGATTCGATATGACGAGCAAGAAAGTGGATCGCATGGT 1397  
 Db 302 CAGGCTTGGGAGGATTTATCTTGATTGCTATATGCAAGAAGCAAAAGTGGATCGCATGGT 361  
 QY 1398 TATCTGCCACAGTTTGAGGAGTACTTGGAGACGGGAAAGTTAGCTTCTGCTCATGCCCA 1457  
 Db 362 TATCTGCCAACGCTTCAGGATATCTTGGAGAACGGGAAAGTTAGCTTCTGCGCATCGCTG 421  
 QY 1458 TGGCACTCAACCCCATCTGACGTTGGACATCCCTTCCCTGATCACATCTCTCAAGGAA 1517  
 Db 422 TCGGCGTTGCAACCCCATGCTGAGATGGACATCCCTTCCCTCTCATCTCTCAAGGAA 481

RESULT 14  
 CF477562 740 bp mRNA linear EST 08-SEP-2003  
 LOCUS R1W3\_8\_G10\_g1\_A022 Well-watered loblolly pine roots W3 Pinus taeda cDNA clone R1W3\_8\_G10\_A022 5', mRNA sequence.  
 DEFINITION CF477562  
 ACCESSION CF477562.1 GI:34506431  
 VERSION CF477562  
 KEYWORDS EST.  
 SOURCE Pinus taeda (loblolly pine)  
 ORGANISM Pinus taeda

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 740)  
 AUTHORS Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and Neale, D.  
 TITLE An EST database from well-watered loblolly pine (Pinus taeda) roots Unpublished (2003)  
 JOURNAL Contact: Cordonnier-Pratt MM  
 COMMENT Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
 Seq primer: JENREV (CAGAAACAGCTATGACC).

FEATURES  
 source Location/Qualifiers  
 1..740  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="CCIONES"  
 /db\_xref="taxon:3352"  
 /clone="R1W3\_8\_G10\_A022"  
 /lab\_host="DH10B-T1 phage-resistant E. coli"  
 /clone\_lib="Well-watered loblolly pine roots W3"  
 /note="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN  
 Query Match 18.6%; Score 375; DB 14; Length 740;  
 Best Local Similarity 76.1%; Pred. No. 1.6e-56;

```
Matches 462; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
QY 1307 TCAAGGCGGACACAGCTCAACTATGCAAGACAGGCTGGGAGGCGTCTTTTGATTCGTA 1366
Db |||||
740 TCAAGGCTGGGATACAGTCAGTCATGCTCGAAATCTTTGGGAGGCTTTTATTTGGTGCTTA 681
QY 1367 TATGACAGAACAAAGTGGATCGCCACTGGTATCTGCCAGGTTTGGAGAGTACTTGGGA 1426
Db |||||
680 TATACAAAGAACCAAGTGGATTTCCAGTGGTATCTTCCAGTTTCGACGAGTACTCTCGA 621
QY 1427 GAAACGGGAAGTTAGCTCTGCTCATCGCCATGCGCACTGCAACCCATTCTTGACGTTGGA 1486
Db |||||
620 GAATGGGAGGTCAGCTTCGGCTCTCGCATACCAAGCTCGAACCCATGCTGACTTTGGG 561
QY 1487 CATCCCTTTCTGATCATCTCTCAAGGAAGTTGACTTCCCATCGAAGCTCAATGACTT 1546
Db |||||
560 GTTTCCTCTCCGCTCGAATCTCGAGGAATTGACTTTCCATCGAAATTCATGATTT 501
QY 1547 GATATGTATCATCTTCTGATTAAGAGGTGATACAGGCTGCTACAGGCGACAGGCGCG 1606
Db |||||
500 GATATGTGCCATCTCTCGACTGAAAGGTGACATCAATGCTACAGGCTGACAGGCGCG 441
QY 1607 TGGAGAAAGCTTCGTCTATATCATCTTATATGAAGACAACTCTGTGATTAACGGAAGA 1666
Db |||||
440 TGGAGAAAGCTTCGCGCGTATGCTGTATATGAAGACCACTCTCGAATTAACAGAGA 381
QY 1667 AGATGCTCTGAATCATATCAATCTTCATGATCAGGAGCGCAATCAGAGAAATTAATTTGGGA 1726
Db |||||
380 AGATGCTCTCAATCAAGTCAATGCTATGCTGATAAATTAACCAAGGAAGTGAATTTGGA 321
QY 1727 GCTTCTAAGCAGACACAGCTGTTCCCATCACTTCCGAAGAACACGCAATTTGACATAAG 1786
Db |||||
320 GTTACTTTAGACCCGACAGCGTGTTCCTCATCTTTACAAAGAGTGTCTTTTGACATTTG 261
QY 1787 CAGAGTTTGGCATACCGTTACAGATACCGAGTACAGCTTGTGCAAGTTTGAAC 1846
Db |||||
260 CAGAGTTTTCATTACGGTTACAAATACAGAGAGGCTTCAGTTGTCCAGTATTTGAAT 201
QY 1847 AAAGAGTTTGGTATGAGAACCGCTGATGAACCTGTGCTTTGTTAAACACACTTCAAATC 1906
Db |||||
200 AAAGAAATTTGTAACAGAACCGTGTGTGAACCTGTGCTTTGAGCCACACATGAAATG 141
QY 1907 TACAATA 1913
Db |||||
140 TACAATA 134
```

```
RESULT 15
AW287756 539 bp mRNA linear EST 07-JAN-2000
LOCUS EST0004 Sitka spruce drill-wounded bark Picea sitchensis cDNA clone
DEFINITION 25-1-3 5' similar to mono-terpene synthase, mRNA sequence.
ACCESSION AW287756
VERSION AW287756.1 GI:6681768
KEYWORDS EST.
SOURCE Picea sitchensis (Sitka spruce)
ORGANISM Picea sitchensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
Wang, S.X., Hunter, W. and Plant, A.L.
AUTHORS Isolation of terpene synthase gene-specific probes from Sitka
TITLE spruce and induction of gene expression by simulated white pine
weevil damage
JOURNAL Unpublished (2000)
COMMENT Contact: Wang SX
Department of Biological Sciences
Simon Fraser University
8888 University Drive, Burnaby, BC V5A1S6, Canada
Tel: 604 584 8870
Fax: 604 584 8873
Email: sxwang@bigfoot.com
Clone sequence of a RT-PCR product from the mRNA of drill-wounded
```

```
Bark tissue
PCR Primers
FORWARD: Mult-F10
BACKWARD: Mult-R18
Insert Length: 539 Std Error: 0.00
Seq primer: M13 Reverse and M13 Forward.
High quality sequence stop: 539.
FEATURES
Location/Qualifiers
source 1..539
/organism="Picea sitchensis"
/mol_type="mRNA"
/db_xref="taxon:3332"
/clone="25-1-3"
/tissue type="bark"
/clone_lib="Sitka spruce drill-wounded bark"
/note="mRNA isolated from drill-wounded bark tissues;
RT-PCR product"
```

## ORIGIN

```
Query Match 17.5%; Score 352.4; DB 10; Length 539;
Best Local Similarity 79.1%; Pred. No. 1.8e-52;
Matches 419; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
QY 1135 TTCTTGACGACATGTAAGACGCTCTTCGSCACAGTAGACGAGCTGGAACCTCTTCACAGCA 1194
Db 1 TGCTCGAGGACCTGTACGACACATTCGGAACAATCGACGAAATCGAATCTTTCACAGAAG 60
QY 1195 CAATTAAAGAGATGGATCCGTCGCGATGGAATGCTTCCAGAAATATATGAAAGAGTGT 1254
Db 61 CAGTCAGAGATGGATCCGTCGAGACAGAGAGCTTCCAGACTATATGAAAGAGTGT 120
QY 1255 ACATGATGTTTATCACACCGTAAATGAAATGGCTCGAGTCGCGAGAGAGGCTCAAGGCC 1314
Db 121 ACATGGTACTCTACGAAGCCCTAACTGAAATGGCTCAAGAGCGGAGAGAAACACAAGGCC 180
QY 1315 GAGACAGCTCAACTATGCAAGACAGGCTTGGGAGGCGTGTGATTCGTATATGCAAG 1374
Db 181 GAGACAGCTCAACTATGCTCGAAAGGCTTGGGAGATTTATCTTGATTCGTATATTCAG 240
QY 1375 AAGCAAGTGGATCGCCACTGTTATCTGCCACGTTTGAGGAGTACTTGGAGAACGGA 1434
Db 241 AAGCAAGTGGATCGCCAGTGGTATCTGCCAACATTTTCAGGAGTACTTTCGAGAACGGA 300
QY 1435 AAGTTAGTCTGCTCATCGCCCATGCGCACTGCAACCCCATTTCTGACGTTGGAACATCCCT 1494
Db 301 AAATTAGTCTGCTTATCGCGCAGCGGCAATTGACACCCCATCTCATTGGACGTACCGC 360
QY 1495 TTCCTGTATCATCTCAAGGAAGTTGACTTCCCATCGAAGCTCAATGACTTGATATGTA 1554
Db 361 TTCCTGATATCATCTTGAAGGGAATTTGATTTCCATCGAGATTCATATGATTTGGCATCTT 420
QY 1555 TCATCCTTCGATTAAGAGGTGATACAGGTCGCTCAAGGCGACAGAGGCGCGCTGGAGAAG 1614
Db 421 CCTTCTTCGACTAAGAGGTGACACAGCTGCTACAGGCGGATAGGGCCGCTGGAGAAG 480
QY 1615 AAGCTTCTGTATATCATGTTTATATGAAGAAATCTCTGGATTAACGGA 1664
Db 481 AAGCTTCGTGATATCTTGTACATGAAGGACCAACCCCACTGCTACATGAA 530
```

Search completed: July 8, 2004, 16:55:39  
Job time : 6172.26 secs

**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2004, 12:09:59 ; Search time 1054.29 Seconds  
(without alignments)  
9305.868 Million cell updates/sec

Title: US-10-025-145A-64  
Perfect score: 2013  
Sequence: 1 ttttgagtgctctttatc.....aaaaaaaaaaaaaaaaaaaaa 2013

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3183909 seqs, 2436941669 residues

Total number of hits satisfying chosen parameters: 6367818

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2013	100.0	2013	15	US-10-025-145A-64
2	1306.8	64.9	2018	9	US-09-887-586A-19
3	1306.8	64.9	2018	9	US-09-903-012-19
4	1306.8	64.9	2018	10	US-09-900-797-19
5	1306.8	64.9	2018	13	US-09-893-820-19
6	1306.8	64.9	2018	14	US-10-041-007-21
7	1306.8	64.9	2018	15	US-10-025-145A-3
8	1092.4	54.3	1890	15	US-10-025-145A-77
9	1072.4	53.3	2186	15	US-10-025-145A-66
10	1071.6	53.2	2196	9	US-09-887-586A-29
11	1071.6	53.2	2196	9	US-09-903-012-29
12	1071.6	53.2	2196	10	US-09-900-797-29
13	1071.6	53.2	2196	13	US-09-893-820-29
14	1071.6	53.2	2196	14	US-10-041-007-25

15 1071.6 53.2 2196 15 US-10-025-145A-1  
16 1071.6 53.2 2205 15 US-10-025-145A-31  
17 929.4 46.2 2429 14 US-10-041-007-27  
18 929.4 46.2 2429 15 US-10-025-145A-68  
19 925.4 46.0 2089 9 US-09-887-586A-57  
20 925.4 46.0 2089 9 US-09-903-012-57  
21 925.4 46.0 2089 10 US-09-900-797-57  
22 925.4 46.0 2089 13 US-09-893-820-57  
23 925.4 46.0 2089 14 US-10-041-007-23  
24 925.4 46.0 2089 15 US-10-025-145A-5  
25 471 23.4 696 15 US-10-025-145A-70  
26 445.8 22.1 1865 9 US-09-887-586A-47  
27 445.8 22.1 1865 9 US-09-903-012-47  
28 445.8 22.1 1865 10 US-09-900-797-47  
29 445.8 22.1 1865 13 US-09-893-820-47  
30 445.8 22.1 1865 14 US-10-041-007-17  
31 429.6 21.3 1967 15 US-10-025-145A-17  
32 427 21.2 2700 9 US-09-887-586A-43  
33 427 21.2 2700 9 US-09-903-012-43  
34 427 21.2 2700 10 US-09-900-797-43  
35 427 21.2 2700 12 US-10-041-018-363  
36 427 21.2 2700 13 US-09-893-820-43  
37 391.6 19.5 2424 9 US-09-887-586A-45  
38 391.6 19.5 2424 9 US-09-903-012-45  
39 391.6 19.5 2424 10 US-09-900-797-45  
40 391.6 19.5 2424 13 US-09-893-820-45  
41 391.6 19.5 2424 14 US-10-041-007-14  
42 391.6 19.5 2424 15 US-10-025-145A-15  
43 391.6 19.5 2528 14 US-10-041-007-13  
44 390.2 19.4 1785 9 US-09-887-586A-49  
45 390.2 19.4 1785 9 US-09-903-012-49

## ALIGNMENTS

## RESULT 1

US-10-025-145A-64  
; Sequence 64, Application US/10025145A  
; Publication No. US20030175861A1  
; GENERAL INFORMATION:  
; APPLICANT: Croteau, Rodney B.  
; APPLICANT: Bohlmann, Joerg  
; APPLICANT: Steele, Christopher L.  
; APPLICANT: Phillips, Michael A.  
; TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)  
; FILE REFERENCE: WSU118414  
; CURRENT APPLICATION NUMBER: US/10/025,145A  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: US 09/360,545  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: PCT/US98/14528  
; PRIOR FILING DATE: 1998-07-10  
; PRIOR APPLICATION NUMBER: US 60/052,249  
; PRIOR FILING DATE: 1997-07-11  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 64  
; LENGTH: 2013  
; TYPE: DNA  
; ORGANISM: Abies Grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (36)..(1889)  
; OTHER INFORMATION:  
US-10-025-145A-64

Query Match 100.0%; Score 2013; DB 15; Length 2013;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2013; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TTTTGAGTGCCTTCCTTATCTGATAGCAAGCTGMAATGGCTCTTCTTCTATTAATCGCG 60

1 TTTTGGCGGCTCTTTATCTGATAGCAAGCTGAAATGGCTCTCTTTCTTATTAATCTCCG 60  
61 TGGTTTCAGGTCGCTCCCTCAGTTCTTCTCATGAGATTAAGGCTCTCCGTTAGAACAAATCC 120  
61 TGGTTTCCAGGTCGCTCCCTCAGTTCTTCTCATGAGATTAAGGCTCTCCGTTAGAACAAATCC 120  
121 CAACTCTTGGAAATCTCAGCGCGGGGAAATCCGTCGCGCATTCCTAATAACATGTGTTGA 180  
121 CAACTCTTGGAAATCTCAGCGCGGGGAAATCCGTCGCGCATTCCTAATAACATGTGTTGA 180  
181 CAAAGCTCGCATCTACTGATCTCTGTACAGAGACGCGTGGGCAACTATCAATTCOAACCTGT 240  
181 CAAAGCTCGCATCTACTGATCTCTGTACAGAGACGCGTGGGCAACTATCAATTCOAACCTGT 240  
241 GGAACGATGATTCATACAGTCTCTGATCTCAACGCTTATGAGACACCTGATTAACGGG 300  
241 GGAACGATGATTCATACAGTCTCTGATCTCAACGCTTATGAGACACCTGATTAACGGG 300  
301 AACGTCGTACAGACTTATTTGGGGAAGTAAAGGATATAATGTTCAATTTCAAAGTCGCTGG 360  
301 AACGTCGTACAGACTTATTTGGGGAAGTAAAGGATATAATGTTCAATTTCAAAGTCGCTGG 360  
361 AAGATGGAGCAATGATCTCTTCAACGACTTTTGTGTCGATGAGCTTGAACGTTGG 420  
361 AAGATGGAGCAATGATCTCTTCAACGACTTTTGTGTCGATGAGCTTGAACGTTGG 420  
421 GAATCAGACAGGCAATTCAAAAAGATAAAAAGGCACTCGATTAATGTTAAACAGTTATT 480  
421 GAATCAGACAGGCAATTCAAAAAGATAAAAAGGCACTCGATTAATGTTAAACAGTTATT 480  
481 GGAACCAAAAAGGCAATTTGAATGTTGGGAGGAGAGTGTGTGACTGACCTCAACTCAACCG 540  
481 GGAACCAAAAAGGCAATTTGAATGTTGGGAGGAGAGTGTGTGACTGACCTCAACTCAACCG 540  
541 CCTTGGGGCTTCGAACCTCCGACTACACGATACACTGTGTCTTCAGATGTTTGAACG 600  
541 CCTTGGGGCTTCGAACCTCCGACTACACGATACACTGTGTCTTCAGATGTTTGAACG 600  
601 TTTTAAAGCAAAAATGGCAATTTCTCCACTGCAATATTCAGATAGAGGAGAGA 660  
601 TTTTAAAGCAAAAATGGCAATTTCTCCACTGCAATATTCAGATAGAGGAGAGA 660  
661 TTAGAGCGCTTCAATTTATTCAGGCGCTCCCTCGCTGCTTCCGCGGAGAAAGTTA 720  
661 TTAGAGCGCTTCAATTTATTCAGGCGCTCCCTCGCTGCTTCCGCGGAGAAAGTTA 720  
721 TGGATGAAGCTGAAACATTTCTACAAAATATTTAAGAGAAGCGCTGCAAAAAGATTCCGG 780  
721 TGGATGAAGCTGAAACATTTCTACAAAATATTTAAGAGAAGCGCTGCAAAAAGATTCCGG 780  
781 CATCCAGTATACCTTCACTAGATACGGGACGTTCTGGAATATGTTGGCAGACCAATTT 840  
781 CATCCAGTATACCTTCACTAGATACGGGACGTTCTGGAATATGTTGGCAGACCAATTT 840  
841 TGCCACGCTTTGGAAGCAAGGAATTTACATGAGCGTCTTTGGACAGCACACTAAAAATAAGA 900  
841 TGCCACGCTTTGGAAGCAAGGAATTTACATGAGCGTCTTTGGACAGCACACTAAAAATAAGA 900  
901 ACGCGCGGAGAAACTTTTGAACCTTCAAAAATTTGGAATTCATATATTTTCACTCCTTAC 960  
901 ACGCGCGGAGAAACTTTTGAACCTTCAAAAATTTGGAATTCATATATTTTCACTCCTTAC 960  
961 AAGAGAGAGAGTTAAAAATGTTTCCGATGTTGGAAGAGACTCGGGTCTCCTGAGATGA 1020  
961 AAGAGAGAGAGTTAAAAATGTTTCCGATGTTGGAAGAGACTCGGGTCTCCTGAGATGA 1020  
1021 CCTTCTGTGACATCTGTCAGTGAATTAACAGCTTTTGGCTTCTGCTGATTTGGCTGAGC 1080  
1021 CCTTCTGTGACATCTGTCAGTGAATTAACAGCTTTTGGCTTCTGCTGATTTGGCTGAGC 1080  
1081 CTCACATCTGGAATTCAGACTCGGCTTTTACAGAGTGTCTCATCTTATCACGCTTCTTG 1140  
1081 CTCACATCTGGAATTCAGACTCGGCTTTTACAGAGTGTCTCATCTTATCACGCTTCTTG 1140

1141 ACGCATGTACGACGCTTTCCGCAAGTAGACGAGCTGAACTCTTTCACAGCGCAATTA 1200  
1141 ACGCATGTACGACGCTTTCCGCAAGTAGACGAGCTGAACTCTTTCACAGCGCAATTA 1200  
1201 AGAGATGGGATCCGTCGCGATGGAATGCTTCCAGATATATGAAAGGAGTGTACATGA 1260  
1201 AGAGATGGGATCCGTCGCGATGGAATGCTTCCAGATATATGAAAGGAGTGTACATGA 1260  
1261 TGGTTTATCACACCGTAAATGAAATGGCTCGAGTGCAGAGAGGCTCAAGGCCGAGACA 1320  
1261 TGGTTTATCACACCGTAAATGAAATGGCTCGAGTGCAGAGAGGCTCAAGGCCGAGACA 1320  
1321 CGCTCAACTATGCAACACAGGCTTTGGAGCGGTGTTTGTATCTGATATATGACAGAGCAA 1380  
1321 CGCTCAACTATGCAACACAGGCTTTGGAGCGGTGTTTGTATCTGATATATGACAGAGCAA 1380  
1381 AGTGGATCGCACTGGTTATTCGCCCAACGCTTTCAGAGTACTTGGAGAACGGGAAAGTTA 1440  
1381 AGTGGATCGCACTGGTTATTCGCCCAACGCTTTCAGAGTACTTGGAGAACGGGAAAGTTA 1440  
1441 GCTCTGCTCATCGCCCATCGCACTGCAACCCCACTCTGAGCTTGGACATCCCTTTCCCTG 1500  
1441 GCTCTGCTCATCGCCCATCGCACTGCAACCCCACTCTGAGCTTGGACATCCCTTTCCCTG 1500  
1501 ATCACATCTCTCAAGGAAGTTGACTTCCCATCGAAGCTCAATGACTTGTATGATCATCC 1560  
1501 ATCACATCTCTCAAGGAAGTTGACTTCCCATCGAAGCTCAATGACTTGTATGATCATCC 1560  
1561 TTCCGATTAAGAGTGATACACGCTGCTACAGGACACAGGCGCCGCTGGAGAGAAAGCTT 1620  
1561 TTCCGATTAAGAGTGATACACGCTGCTACAGGACACAGGCGCCGCTGGAGAGAAAGCTT 1620  
1621 CGTCTATCATCTTATATGAAAGACAATCTCGATTAAACGAGAGAGATGCTCTGAATC 1680  
1621 CGTCTATCATCTTATATGAAAGACAATCTCGATTAAACGAGAGAGATGCTCTGAATC 1680  
1681 ATATCAACTTTCATGATCAGGACGCAATCAGAGAAATTAATTTGGAGCTTCTAAAGCCAG 1740  
1681 ATATCAACTTTCATGATCAGGACGCAATCAGAGAAATTAATTTGGAGCTTCTAAAGCCAG 1740  
1741 ACAAGAGTGTTCCTCATCTTCCAGAAACACGCTTTCAGATAGAGAGTTCGATC 1800  
1741 ACAAGAGTGTTCCTCATCTTCCAGAAACACGCTTTCAGATAGAGAGTTCGATC 1800  
1801 ACGGTTACAGATACCGAGATGCTTACAGCTTTTCCCAACGCTTGAACAGAGTTCGTGA 1860  
1801 ACGGTTACAGATACCGAGATGCTTACAGCTTTTCCCAACGCTTGAACAGAGTTCGTGA 1860  
1861 TGAGAACCGTTCATTTGAACCTGTGCTTTTGAACACACTTCAATCTCAATATTAATCTG 1920  
1861 TGAGAACCGTTCATTTGAACCTGTGCTTTTGAACACACTTCAATCTCAATATTAATCTG 1920  
1921 AGGATGCCCTATCGGTTGATATAGGCGACACAAAATAAATATGTTGCTGTAGTAAAGC 1980  
1921 AGGATGCCCTATCGGTTGATATAGGCGACACAAAATAAATATGTTGCTGTAGTAAAGC 1980  
1981 TGTAAATTTGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2013  
1981 TGTAAATTTGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2013

RESULT 2  
US-09-887-586A-19  
; Sequence 19, Application US/09887586A  
; Patent No. US20020094556A1  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. US20020094556A1, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001

```
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6)...(1889)
; OTHER INFORMATION: pinene synthase
US-09-887-586A-19

Query Match      64.9%; Score 1306.8; DB 9; Length 2018;
Best Local Similarity 81.0%; Pred. No. 0;
Matches 1595; Conservative 0; Mismatches 342; Indels 33; Gaps 5;

QY 68 CAGGTCGTCGCTCAGTCTTCTCATGAGATTAAGGCTCTCCGTAGAACCAATCCCAACTCT 127
DB 53 CAATCGTGTGATCAGTCTTCTACCATGAGCTTAAGGCTCTCTCTAGAACCAATCCCAACTCT 112

QY 128 TGGAACTCTGACGGCCGGGGAATCCGTCGCGCATTTCCCATTAACATGTGTTTGACAAGCGT 187
DB 113 AGGAATGAGTAGGCGAGGGAATCTATCACTCTTCCATCAGCATGAGCTCTACACCGT 172

QY 188 CGCATCTACTGATTCGTACAGAGACGGTGGGCAACTATCATTCCAACCTGTGGGACGA 247
DB 173 TGTAAACCGATGATGTGTACGAAGACGCATGGCGATTTCCATTTCAACCTCTGGGACGA 232

QY 248 TGATTTCTATACAGTCTCTGATCTCAACGCCCTTATGGAGCACCTGATACCGGGAACGTGC 307
DB 233 TGATGTCATACAGTCTCTTACCACCGGCTTATGAGGAAAAATCGTACCTGGAGCGTGC 289

QY 308 TGACAGACTTATTTGGGGAAGTAAAGGATATATGTTTCAATTTCAAGTCGCTGGAAGATGG 367
DB 290 TGAGAACTGATCGGGGAAGTAAAG--AACATGTTCAATTCGATGTCATTAGAGATGG 346

QY 368 AG-----GCAATGATCTCTTCAACGACTTTTGCTGGTGCATGAGCTTGA 412
DB 347 AGAGTTAATGATCGCTCAATGATCTCATTTCAACGCCCTTTGGATTGTCGACAGCCTTGA 406

QY 413 ACGTTTGGGAATCGACAGGCAATTTCAAAAAGAGATAAAACGGCACTCGATTATGTTAA 472
DB 407 ACGTTTGGGATCCATAGACATTTCAAAGATGAGATAAAATCGGCGCTTGAATTATGTTTA 466

QY 473 CAGTTATTGGAACGAAAAGCAATTTGGATGTGGGAGGAGAGTGTGTGACTGACCTCAA 532
DB 467 CAGTTATTGGGCGAAAAATGSCATCGGATCGGGAGGAGAGTGTGTGACTGATCTGAA 526

QY 533 CTCACCGCCTTGGGCTTGAACTCTCCGACTACACGGATACACTGTGTCTCAGATGT 592
DB 527 CTCACCTGCTGGGCTTTCGAAACCTACGACTACACGGATACACCGGCTGTCTCAGATGT 586

QY 593 TTTGAACGTTTTTAAAGACAAAAATGGGCAATTTTCTCCACTCCCAATATTCAGATAGA 652
DB 587 TTTCAAAGCTTTCAAAGGCCAAAATGGGCAGTTCCTGCTCTGAAAATAATTCAGACAGA 646

QY 653 GGGAGAGATTAGGCGTTCCTCAATTTATTCAGGCGCTCCCTCGTCGCGCTTTCCCGGGA 712
DB 647 TGAAGAGATCAGAGCGTTCGAATTTATTCGGGCTCCCTCATTCGCTTTCCAGGGA 706

QY 713 GAAAGTTATGATGAGCTGAAACATCTCTCAAAAATATTTAAGAGAGAGCCCTGCAAAA 772
DB 707 GAAAATATGATGAGGCTGAAATCTCTCTACCAAAATATTTAAAGAGAGCCCTGCAAAA 766

QY 773 GATTCCGSCATCCAGTATACATTCTACTAGATACGGGACGTTCTGGGAATATGGTTGCA 832
```

```
DB 767 GATTCGGGTCTCCAGT---CTTTCGAGAGATCGGGACGTTTGGAAATATGGTTGGCA 823
QY 833 CACCAATTTGGCAGCGCTTGGAGGCAAGCAATTTACATGACGCTTTTGGACAGCACTTAA 892
DB 824 CACATATTTGGCGGATTTGGAAAGCAAGAAATTTACATCCAAGTCTTTGGACAGGACACTGA 883
QY 893 AAATAAGAAACGC-----CGCCGAGAAACTTTTAGAACTTTGCAAAATTTGGAATTCAA 943
DB 884 GAACACGAAGTCATATGTGAAGAGCAAAAACCTTTTAGAACTCGCAAAATTTGGAGTTCAA 943
QY 944 TATATTTCACTCTTTACAGAGAGAGATTAAAACATGTTTCCCGATGGTGAAGACTC 1003
DB 944 CATCTTTCAATCTTTACAAAAGAGGAGTGTAGAAAGTCTGGTTCAGATGGTGAAGAAATC 1003
QY 1004 GGGTTCTCTGAGATGACCTTCTGTCGACATCGTCACGTCGNAATACACTAGCTTTGGCTTC 1063
DB 1004 GGGTTTTCTGAGATGACCTTCTGCGGACATCGTCACGTCGNAATACACTAGCTTTGGCTTC 1063
QY 1064 CTGCATTCGGTTCCGAGCCCTCAACATTTCTGGATTTCAGACTCGGCTTTTACCAAGATCTCA 1123
DB 1064 CTGCATTCGGTTCCGAGCCCTCAACATTTCTGGATTTCAGACTCGGCTTTTCCAGACGTTCA 1123
QY 1124 TCTATCAGCGTTCTTGAGACATGTGACAGCTCTTTCGGGACAGTAGACGAGCTGGAAT 1183
DB 1124 TCTTATCACGGTTCTTGACGATATGTACGACACCTTTCGGGACAGTAGACGAGCTGGAAT 1183
QY 1184 CTTTCACAGCGCAATTTAAGAGATGGGATCCGTCGCGATGAAATGCCCTTCCAGATATAT 1243
DB 1184 CTTTCACAGCGCAATTTAAGAGATGGGATCCGTCCTCGATAGATTGCCCTTCCAGATATAT 1243
QY 1244 GAAAGGAGTGTACATGATGTTTATCACACCGTAAATGAAATGGCTCGAGTGGCAGAGAA 1303
DB 1244 GAAAGGAGTGTACATGATGTTTATCACACCGTAAATGAAATGGCTCGAGGAGCAGAGGA 1303
QY 1304 GGCTCAAGCGCGAGACAGCTCACTATGTAAGACAGCGCTTGGGAGGCGTGTGTTGATTC 1363
DB 1304 GGCTCAAGCGCGAGATACGCTCACATATGCTCGGGAAGCTTTGGGAGGCTTATATTGATTC 1363
QY 1364 GTATATGACGAAGCAAGTGSATCGCCACTGTGTTATCTGCCACGCTTTGGAGGACTTT 1423
DB 1364 GTATATGCAAGAGCAAGTGGATCGCCACTGTGTTTACCTGCCCTCTTTGATGACTATA 1423
QY 1424 GGAGAACGGGAAAGTTAGCTCTGCTCATCGGCCCATGCGCACTGCAACCCATTTCTACGTT 1483
DB 1424 CGAATATGGGAAAGTTAGCTGTGTTGTCATCGCATATCCGCATTTCAACCCATTTCTGACAT 1483
QY 1484 GGACATCCCTTTCTGATACATCTCTCAAGAAAGTTGACTTCCCATCGAAGCTCAATGA 1543
DB 1484 GGACATCCCTTTCTGATCATATCTCTCAAGAAAGTTGACTTCCCATCAAAGCTTAAAGGA 1543
QY 1544 CTTGATATGATATCATCTCTTCGATTAAAGAGGTATACCGGTGCTACAGGCGCAGAGGCG 1603
DB 1544 CTTGGCATGTCCATCTCTCGATTACAGGTTGATACGGGTGCTTACAGGCGGACAGGCG 1603
QY 1604 CCGTGGAGAGAAAGCTTCTGCTATATCATGTTATATGAAGACAAATCTCTGATTAAACGGA 1663
DB 1604 TCGTGGAGAGAAAGCTTCTCTATATCATGTTATATGAAGACAAATCTCTGAGTATCAGA 1663
QY 1664 AGAAGATGCTCTGATATCATATCACTTCATGATCAGGAGCGCAATCAGAGAAATTAATTG 1723
DB 1664 GGAAGATGCTCTCGATCATATCAACGCGCATGATCAGTGACGTAAATAAAGGATTTAAATTG 1723
QY 1724 GGAGCTTTCAAAGCCAGACACAGTGTTCCTCATCACTTTCCAAGAAAACACGCTTTTGACAT 1783
DB 1724 GGAATTTCTCAAACAGACATCAATGTTCCCATCTCGGCGAAGAAACATGCTTTTGACAT 1783
QY 1784 AAGCAGAGTTTGGCATCAGGTTTACAGATACCGAGATGGCTTACAGCTTTTGCACAGTTGA 1843
DB 1784 CGCCAGAGCTTTTCCATTTACGGCTACAAATACCGAGACGGCTTACAGCGTTGCGCAAGTTGA 1843
QY 1844 AACAAAGAGTTTGGTGTATGAGAACCGCTCATGTAACCTGTGCTGCTTTGTAACAACTTCAA 1903
```

Db 1844 AACGAAGAGTTGGTCACGAGAACCCCTCCTGTAATCTGTGCTTTGTAGCAACAGCTCAA 1903  
 Qy 1904 ATCTACAATTAATACGAGATGCGCTATGGGTGTATATAGGCGACACAAAAATAATAT 1963  
 Db 1904 ATCTATGCCCTATGCTATGTCGGGTGTAATAATATATGTGAAGGTAGCCGTGGATGTAGA 1963  
 Qy 1964 GGTGTGTAGTAAAGCTGTAAATTTATGAAAAAATGAAAAAATGAAAAAATGAAAAA 2013  
 Db 1964 GGTAAAGTTGTTATTAATTAATAAAGTTGTAATTTAAAAAATGAAAAAATGAAAAA 2013

RESULT 3  
 US-09-903-012-19  
 ; Sequence 19 Application US/09903012  
 ; Patent No. US20020094557A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chappell, Joseph  
 ; APPLICANT: No. US20020094557A1, Joseph P.  
 ; APPLICANT: Starks, Courtney M.  
 ; APPLICANT: Manna, Kathleen R.  
 ; TITLE OF INVENTION: SYNTHASES  
 ; FILE REFERENCE: 07678-025001  
 ; CURRENT APPLICATION NUMBER: US/09/903,012  
 ; CURRENT FILING DATE: 2001-07-11  
 ; PRIOR APPLICATION NUMBER: 09/398,395  
 ; PRIOR FILING DATE: 1999-09-17  
 ; PRIOR APPLICATION NUMBER: 60/100,993  
 ; PRIOR FILING DATE: 1998-09-18  
 ; PRIOR APPLICATION NUMBER: 60/130,628  
 ; PRIOR FILING DATE: 1999-04-22  
 ; PRIOR APPLICATION NUMBER: 60/150,262  
 ; PRIOR FILING DATE: 1999-08-23  
 ; NUMBER OF SEQ ID NOS: 58  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 19  
 ; LENGTH: 2018  
 ; TYPE: DNA  
 ; ORGANISM: Abies grandis  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (6)...(1889)  
 ; OTHER INFORMATION: pinene synthase  
 US-09-903-012-19

Query Match 64.9%; Score 1306.8; DB 9; Length 2018;  
 Best Local Similarity 81.0%; Pred. No. 0;  
 Matches 1595; Conservative 0; Mismatches 342; Indels 33; Gaps 5;  
 Qy 68 CAGGTGCGTCTCAGTTCCTCATGAGATTAAAGGCTCTCCGTAGAACAAATCCCACTCT 127  
 Db 53 CAAATCGTTGATCAGTTCTACCCATGAGCTTAAGGCTCTCTTAGAACAAATCCAGCTCT 112  
 Qy 128 TGAATCTGAGCGCGGGAATCCGTCGGGCAATCCATAACATGTTGTTGACAGCGT 187  
 Db 113 AGGAATGAGTAGCGAGGGAATCTATCACTCTCTCCATCAGCATGAGCTCTACCCCGT 172  
 Qy 188 CGCATCTACTGATTCTCTACAGAGACGCTGGGCACTATCATTCACCACTGTGGGACGA 247  
 Db 173 TGTAAACGATGATGTTGATGAGACGCAATGGGCGATTCATTCACCACTCTGGGACGA 232  
 Qy 248 TGATTTTCATACAGTCTCTGATCTCAACGCTTATGGAGCACTGATTTACCGGGAACGTGC 307  
 Db 233 TGATGTCATACAGTCT--TTACCAACGGCTTATGAGGAAATTCGATCTGGAGCGTGC 289  
 Qy 308 TGACAGACTTATGGGGAAGTAAAGATATAATGTTCAATTTCAAGTCGGTGGAGATGG 367  
 Db 290 TGAGAACTGATCGGGGAAGTAAAG---AATGTTTCAATTCGATGTCATTAGAAGATGG 346  
 Qy 368 AG-----GCAATGATCTCTTCAACGACTTTTGTGCTGCTGATGACGTTGA 412  
 Db 347 AGACTTAATGAGTCGGCTCAATGATCTCAATTCACGCTTTGGATTGTCGACAGCCTTGA 406  
 Qy 413 ACGTTTGGGAATCGACAGGCAATTTCAAAAAAGATATAAAACGGCACTCGATTATGTATA 472

Db 407 ACCTTTGGGATCCATAGACATTTCAAAGATGAGATAAAATCGGCGCTTGAATATGTTA 466  
 Qy 473 CAGTTATTGGAACGAAAGGATGATGATGCGGAGGAGAGTGTGTGCTACTCACCTCAA 532  
 Db 467 CAGTTATTGGGCGAAATGGAATGCGATCGGATCGGAGGAGAGTGTGTACTATCTGAA 526  
 Qy 533 CTCAACGCGCTTGGGCTTTCGAACTCTCGACTTACACGGATACACTGTGTCTTCAGATGT 592  
 Db 527 CTCAACTGCGTTGGGCTTTCGAACTCTACAGCTTACAGGATACCCGCTGTCTTCAGATGT 586  
 Qy 593 TTTGAAAGCTTTTAAAGACAAATGCGCAATTTTCTCTCACTGCGCAATATTCAGATAGA 652  
 Db 587 TTTCAAAGCTTTCAAAGGCGCAATGCGCAATTTTCTCTCTCTCTCTCTCTCTCTCTCT 646  
 Qy 653 GGGAGAGATTAGAGCGTCTCTCAATTTATTCAGGCGCTCTCTCTCTCTCTCTCTCTCTCT 712  
 Db 647 TGAAGAGATCAGAGCGTCTCTGAATTTATTCGGGCGCTCTCTCTCTCTCTCTCTCTCT 706  
 Qy 713 GAAAGTTATGGATGAAGCTGAACATTTCTCTCAAAATATTTAAGAGAGAGCCCTGCAAAA 772  
 Db 707 GAAATTTATGATGAGCTGAATCTCTCTCAAAATATTTAAGAGAGAGCCCTGCAAAA 766  
 Qy 773 GATTCGGCATCCAGTATTTCTCTAGAGATACGGGAGCTTCTGGAATATGTTGGCA 832  
 Db 767 GATTCGGCTCTCCAGT---CTTTCGGGAGAGATCGGGGAGCTTTTGGAAATATGTTGGCA 823  
 Qy 833 CACCAATTTGCCACGCTTGAAGCAAGGAATACATGAGCTCTTTGGAGAGACACACTAA 892  
 Db 824 CACATATTTGCCCGCATTTGAAGCAAGGAATACATCCAACTCTTTGGAGAGACACTGA 883  
 Qy 893 AAAATAAGAACGC-----CGCCGAGAACTTTTAGAACTTGCAGAAATTTGAAATTTCAA 943  
 Db 884 GAACAGAGTCTATATGTAAGAGAGAAAAAATTTTAGAACTCGCAAAATTTGAGTTCAA 943  
 Qy 944 TATATTTCTCTTACAGAGAGAGAGTTTAAACATGTTTCCGATGTTGGGAAAGCTC 1003  
 Db 944 CATCTTTCAATCTCTTACAAAGAGAGAGTTTAAAGTCTGGTTCAGATGTTGGAAGATC 1003  
 Qy 1004 GGGTCTCTGAGATGACCTTCTGTGACATCTGCTGAGTGAATACCTACGCTTTGGCTTC 1063  
 Db 1004 GGGTTTTCTGAGATGACCTTCTGCCGACATCTGCTGAGTGAATACCTACGCTTTGGCTTC 1063  
 Qy 1064 CTGCAATTTGGCTTCGAGCTCAACATTTCTGATTCAGACTCGGCTTTGCCAAGACGTGCA 1123  
 Db 1124 TCTTATCAGCGTTCTTGAAGACATGACAGCTTCTCGGACACAGTACAGAGCTGGAAC 1183  
 Qy 1124 TCTTATCAGCGTTCTTGAAGACATGACAGCTTCTCGGACACAGTACAGAGCTGGAAC 1183  
 Db 1184 CTTTACAGCGCAATTAAGAGATGGATCGTCCGATGGAATGCTTCCAGAAATATAT 1243  
 Qy 1184 CTTTACAGCGCAATTAAGAGATGGATCGTCCGATGGAATGCTTCCAGAAATATAT 1243  
 Qy 1244 GAAAGGAGTGTACATGATGTTTATCACCGTAAATGAAATGCTTCGATGTCGAGAGAA 1303  
 Db 1244 GAAAGGAGTGTACATGATGTTTATCACCGTAAATGAAATGCTTCGATGTCGAGAGAA 1303  
 Qy 1304 GGTCTAAGGCGGAGACACGCTCAACTATGCAAGACAGGCTTGGGAGCGTGTGTTGATTC 1363  
 Db 1304 GGTCTAAGGCGGAGACACGCTCAACTATGCTCGGAGAGCTTGGAGGCTTATATGATTC 1363  
 Qy 1364 GTATATGCAAGAGCAAAAGTGGATCGCACTGTTATCTCCCAACGCTTTTGGAGGATCTT 1423  
 Db 1364 GTATATGCAAGAGCAAAAGTGGATCGCACTGTTATCTCCCAACGCTTTTGGAGTACTA 1423  
 Qy 1424 GGAAGAGCGGAAAGTTAGCTCTGCTCATCGCCCATGCGCACTGCAACCCCATTTCTGAGCTT 1483  
 Db 1424 CGAAGATGGGAAAGTTAGCTCTGCTCATCGCCCATGCGCACTGCAACCCCATTTCTGAGCTT 1483  
 Qy 1484 GGCATCTCCCTTCTCTGATCAGATCTTCAAGGAGTTGACTTCCATCTGAGCTCAATGA 1543

Db 1484 GGACATCCCTTTCTCTGATCATATCCTCAAGGAAGTTGACTTCCCATCAAGCTTAAACA 1543  
Qy 1544 CTTGATATGATATCATCTCTTGGATTAGAGGTGATACACGGTGTCTCAAGGCACAGAGGC 1603  
Db 1544 CTTGGCATGTGCATCTCTTGGATTAGAGGTGATACGGTGTCTCAAGGCACAGAGGC 1603  
Qy 1604 CGCTGGAGAGAGCTTCTCTATATCATCTTATATGAAAGCAATCTCGGATTAAACGA 1663  
Db 1604 TCGTGGAGAGAGCTTCTCTATATCATCTTATATGAAAGCAATCTCGGAGTATCAGA 1663  
Qy 1664 AGAAGATGCTCTGAATCATATCAATCTCATGATCAGGACGCAATCAGAGAAATTAATG 1723  
Db 1664 GGAAGATGCTCTGATCATATCAAGCCATGATCAGTACGTATCAAGGATTAATG 1723  
Qy 1724 GGAGCTTCTAAAGCAGACAAACAGTGTTCCTCATCTTCCATCACTTCCAGAAACAGCAT 1783  
Db 1724 GGAATCTCTCAACACAGACATCAATGTTCCTCATCTTCCATCACTTCCAGAAACAGCAT 1783  
Qy 1784 AAGCAGAGTTTGGCATCAGGTTACAGATACCGAGATGGCTACAGCTTGGCAAGTTGA 1843  
Db 1784 CGCAGAGCTTTCATTAACGCTTACAAATACCGAGACGCTACAGCTTGGCAAGTTGA 1843  
Qy 1844 AACAAAGAGTTTGGTATGAGAACCGCTCATTAACCTGTGCTTGTAAACACACTTCAA 1903  
Db 1844 AACGAAGATTGTGTCACGAGAACCTCTTGAATCTGTGCTTGTAGCAACAGCTCAA 1903  
Qy 1904 ATCTACAATATTAACAGAGATGACCTATGAGGTGATATAGGCAACAAAAATAATAT 1963  
Db 1904 ATCTATGCCCTATGCTATGTGCGGTAAATATATGTGAAGGTAGCGCTTGGATGTAGA 1963  
Qy 1964 GTTGTGTTAGTAAGCTGAATTTATGAATAAAAAAATAAAAAA 2013  
Db 1964 GGATAAGTTGTTAATAATTAAGTTGAATTTAAAAAATAAAAAA 2013

## RESULT 4

US-09-900-797-19  
; Sequence 19, Application US/09900797  
; Publication No. US20030087406A1  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. US20030087406A1, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/900,797  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: US/09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 19  
; LENGTH: 2018  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (6)...(1889)  
; OTHER INFORMATION: pinene synthase  
US-09-900-797-19

Query Match 64.9%; Score 1306.8; DB 10; Length 2018;  
Best Local Similarity 81.0%; Pred. No. 0;  
Matches 1595; Conservative 0; Mismatches 342; Indels 33; Gaps 5;  
Qy 68 CAGGTCTGCTCAGTCTCTCTCATGAGATTAAAGCTCTCGGTAGAACAAATCCCACTCT 127  
Db 53 CAAATCGTTGATCAGTCTTCAACCATGAGTTAAGGCTCTCTAGAACAAATCCAGCTCT 112

Qy 128 TGGATCTGCAGCGCGGAAATCCGTGCGCATTCATTAACATGTGTTTGCACAGGT 187  
Db 113 AGGAATGATAGGCGAGGAAATCTATCACTCTTCCATCAGCATGAGCTCTACACCGT 172  
Qy 188 CGCATCTACTGATCTGTACAGAGCGGTGGCAACTATCATCTCAACTCTGTGGAGCA 247  
Db 173 TGTAAACCATGATGGTGTACGAAGACGATGGCGATTTCATTCCTCAACCTCTGGAGCA 232  
Qy 248 TGATTTATACAGTCTCTGATCTCAACGCTTATGAGCACTGATTAACCGGAAAGTGC 307  
Db 233 TGATGTCATACAGTCT---TTACCAACGCTTATGAGGAAATCGTACCTGGAGCGTGC 289  
Qy 308 TCACAGACTTATGCGGAAAGTAAAGGATATATGTTTCAATTTCAAGTCCGTGGAAGATGG 367  
Db 290 TGAGAACTGATCGGGGAAAGTAAAG---AACATGTTCAATTCGATGTCATTAGAGATGG 346  
Qy 368 AG-----GCAATGATCTCTTCAACGACTTTTGTGCTGCGATGACGTTGA 412  
Db 347 AGAGTTAATGATCGCTCAATGATCTCAATCAACGCTTGGATTGTGACAGCCTTGA 406  
Qy 413 ACCTTTGGGAATCGACAGCATTTCAAAAGAGATATAAAGCGCACTCGATTATGTTAA 472  
Db 407 ACCTTTGGGAATCGACAGCATTTCAAAAGAGATATAAAGCGCGCTTGATTATGTTA 466  
Qy 473 CAGTTATTTGGAACGAAAGGCAATTTGATGTGGAGGAGAGTGTGTTGACTGACCTCAA 532  
Db 467 CAGTTATTTGGCGGAAATGCAATCGATCGGGAGGAGAGTGTGTTACTGATCTGAA 526  
Qy 533 CTCACCGCTTTGGGCTTTCGAACTCTCCGACTACACGGATACCTGTGCTTCAGATGT 592  
Db 527 CTCACCTGCTTGGGCTTTCGAACTCTCCGACTACACGGATACCTGTGCTTCAGATGT 586  
Qy 593 TTTGAACGTTTTTAAAGCAAAATGGCAATTTTCTCCACTGCAATATTTACAGATGA 652  
Db 597 TTTCAAGCTTTCAAGGCCAAATGGCAGTTTTTCTGCTCTGAAATATTTACAGACA 646  
Qy 653 GGGAGAGATTAGCGGCTTCTCAATTTATTCAGGGCTCTCCCTGCTGCGCTTTCCCGGCA 712  
Db 647 TGAAGAGATCAGAGCGCTTCTGAATTTATTCGGGCGCTCTCCCTCATTTGCTTCCAGGGA 706  
Qy 713 GAAAGTTATGATGAAGTGAACATTTCTACAAATATTTAGAGAGCGCTTGCAAAA 772  
Db 707 GAAATATGATGAGCGTGAATCTTCTACCAATATTTAAAGAGCGCTTGCAAAA 766  
Qy 773 GATTCGGCATCCAGTATACCTTCACTAGAGATACGGACCTTCTGGAATATGTTGGCA 832  
Db 767 GATTCGGCTCTCAGT---CTTTGCGAGAGATCGGGACGTTTTTGGAAATATGTTGGCA 823  
Qy 833 CACCAATTTGCCACGCTTGGAGCAAGGAATTACATGAGACGCTTTTGGACAGCACATA 892  
Db 824 CACATATTTGCGCGGATTTGAAGCAAGGAATTACATCCAACTTTTGGACAGGACACTGA 883  
Qy 893 AAATAAGAACGC-----CGCGGAGAACCTTTTAGAATCTGCAAAATTTGGAATCAA 943  
Db 884 GAACACGAAGTCATATGTGAAGCAAAAACTTTTAGAATCTGCAAAATTTGGAATCAA 943  
Qy 944 TATATTTCACTCTTCAAGAGAGAGAGTTAAACATGTTTTCCCGATGTTGGAAAGACTC 1003  
Db 944 CATCTTTCAATCTTCAAAAGAGGAGATTAGAAAGTCTGGTCAGATGTTGGAAAGACTC 1003  
Qy 1004 GGGTTCCTTGAGATGACCTTCTGTGACATCGTACGTGGAATFACTACGCTTTGGCTTC 1063  
Db 1004 GGGTTTTCTTGAGATGACCTTCTGCGGACATCGTACGTGGAATFACTACCTTTGGCTTC 1063  
Qy 1064 CTGCAATTTGCGTTCGAGCTCAACATTTCTGGAATTCAGACTCGGCTTTGCCAAGAGCTGCA 1123  
Db 1064 CTGCAATTTGCGTTCGAGCTCAACATTTCTGGAATTCAGACTCGGCTTTGCCAAGAGCTGCA 1123  
Qy 1124 TCTTATCAGCGTTCTTTCAGCAGCATGTACGAGCTTCTCGGCACAGTACGAGCTGGAAC 1183  
Db 1124 TCTTATCAGCGTTCTTTCAGCAGCATGTACGAGCTTCTCGGCACAGTACGAGCTGGAAC 1183

```

1184 CTTACAGCGCAATTAAGAGATGGATCCGTCGCGATGGAATGCTTCCAGAAATATAT 1243
1184 CTTACAGCGCAATTAAGAGATGGATCCGTCGCGATGGAATGCTTCCAGAAATATAT 1243
1244 GAAAGAGGTGATCATGATGCTTATACACCGGTAAATGCTGAGAGTGGCAGAGAA 1303
1244 GAAAGAGGTGATCATGATGCTTATACACCGGTAAATGCTGAGAGTGGCAGAGAA 1303
1304 GGCCTAAGCCGAGACACCTCACTATGACAGACAGGCTTGGAGCGGTGTTTGATTC 1363
1304 GGCCTAAGCCGAGATACCTCACTATGACAGACAGGCTTGGAGCGGTGTTTGATTC 1363
1364 GTATATGAGAGAGCAAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423
1364 GTATATGAGAGAGCAAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423
1424 GGAGAACGGGAAAGTTAGTCTGCTCATCGCCCATCGGCACTGCAACCCCATTCGAGTT 1483
1424 CGAGAAATGGGAAAGTTAGTCTGCTCATCGCCCATCGGCACTGCAACCCCATTCGACAAT 1483
1484 GGCATCCCTTTCCTGATCAGATCCTCAAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1543
1484 GGCATCCCTTTCCTGATCATATCCTCAAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1543
1544 CTTGATATGATCATCTCTGATTAAGAGAGTGATACACCGGTGCTACAAGGAGCAGAGGGC 1603
1544 CTTGATATGATCATCTCTGATTAAGAGAGTGATACACCGGTGCTACAAGGAGCAGAGGGC 1603
1604 CCGTGAGAGAAAGCTTCTGATCATCACTGATGATGATGATGATGATGATGATGATGATG 1663
1604 TCGTGAGAGAAAGCTTCTGATCATCACTGATGATGATGATGATGATGATGATGATGATG 1663
1664 AGAGATGCTCTGAATCATATCACTGATGATGATGATGATGATGATGATGATGATGATG 1723
1664 GGAAGATGCTCTGATCATCACTGATGATGATGATGATGATGATGATGATGATGATGATG 1723
1724 GGAGCTCTTAAAGCCAGACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1783
1724 GGAATCTTCAACACGACATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1783
1784 AAGCAGAGTTGGATCAGCGTTACAGATACAGATACAGATACAGATACAGATACAGATAC 1843
1784 CGCCAGAGCTTTCATACGCTGCTACAAATACCGAGAGCGGCTACAGCGGCTTCCCAACG 1843
1844 AACAAAGAGTTGCTGATGAGAACCGGTCATGATGATGATGATGATGATGATGATGATGAT 1903
1844 AACAAAGAGTTGCTGATGAGAACCGGTCATGATGATGATGATGATGATGATGATGATGAT 1903
1904 ATCTACAATATTAACAGAGATGCGCTATGCGGTGATATAGGGCAGACACAAAAATAATAT 1963
1904 ATCTATGCCCTATGCTATGCTGCTTAAATATATATGGAAGGTAGCGGTGATGATGATG 1963
1964 GGTGTTGTTAGTAAGCTGTAATTTATGAAAAAATAATGAAAAAATAATGAAAAAATA 2013
1964 GGATAAGTTGTTGTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2013

```

RESULT 5

```

US-09-893-820-19
; Sequence 19, Application US/09893820
; Publication No. US20040053386A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20040053386A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993

```

```

; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6)...(1889)
; OTHER INFORMATION: pinene synthase
US-09-893-820-19

```

```

Query Match      64.9%; Score 1306.8; DB 13; Length 2018;
Best Local Similarity 81.0%; Pred. No. 0;
Matches 1595; Conservative 0; Mismatches 342; Indels 33; Gaps 5;

```

```

QY 68 CAGGTCGCTCGCTCAGTCTCTCTCATGAGATTAAGGCTCTCCGTAGAACCAATCCCACTCT 127
DB 53 CAAATCGTGTATCAGTCTTACCCATGAGCTTAAGGCTCTCTCTAGAACCAATCCCACTCT 112
QY 128 TGAATCTCGAGCGCCGGGAAATCCGTCGCGCATTCATAAACCATGTGTTTGAACGCT 187
DB 113 AGGAATGAGTGGCGAGGAAATCTATCACTCTCCATCAGCATGAGCTCTACCACTGT 172
QY 188 CGCATCTACTGATCTCTGACAGAGACGCTGGGCAATCATTCACACTGTGGGACGA 247
DB 173 TGTAAACCGATGAGTGTGTACGAGACGATGGGCGATTTCCATTCACACTGTGGGACGA 232
QY 248 TGATTTTCATACGCTCTCTGATCTCAACGCTTATGAGACACCTGATACCGGAAACGTGC 307
DB 233 TGATGCTATACGCTCT---TTACCAACGCTTATGAGGAAATCGTACTGTGAGCGTGC 289
QY 308 TGACAGACTTATTTGGGAAAGTAAAGGATATAATGTTCAATTTCAAGTGGTGGGAAAGTG 367
DB 290 TGAGAAACTGATCGGGGAAAGTAAAG---AACATGTTCAATTCGATGTCTATTGAAGATGG 346
QY 368 AG-----GCAATGATCTCTTCAACGACTTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 412
DB 347 AGAGTTTAATGAGTCCGCTCAATGATCTCAATTCACGCGCTTTGGATTGTGACAGCTTGA 406
QY 413 ACCTTTGGGAAATCGACAGGCAATTTCAAAAAAGAGATAAAAAACGCGCATCTGATTTATGTTAA 472
DB 407 ACCTTTGGGAAATCGACAGGCAATTTCAAAAAAGAGATAAAAAACGCGCTTGTGATTTGTTA 466
QY 473 CAGTTATTTGGAACGAAAAAGGCAATTTGATGTTGGGAGGAGAGTGTGTTGACTGACCTCAA 532
DB 467 CAGTTATTTGGGCGGAAAAATGGCATCGGATCGGGAGGAGAGTGTGTTTACTGATCTGAA 526
QY 533 CTCAACCGGCTTGGGCGTTCGAACTCTCCGACTACACGAGATACACTGTGCTGCTCAGATGT 592
DB 527 CTCAACTGCGTGGGCGTTCGAAACCTTACACGATACACGAGATACCGGCTGCTCAGATGT 586
QY 593 TTTGAAACGTTTTTAAAGACAAAAATGGCAATTTTCTCCACTGCGCAATATTTCAGATAGA 652
DB 587 TTTCAAGCTTTTCAAGGCCCAAAATGGGCAAGTTTCTCTGCTCTGAAAAATATTTCAGACAGA 646
QY 653 GGGAGAGATTAGAGGCGTCTCAATTTTATTCAGGCGCTCCCTGCTGCTGCTGCTGCTGCTGCTG 712
DB 647 TGAAGAGATCAGAGGCGTCTGAAATTTTATTCGCGGCGCTCCCTCATTCGCTTTCCAGGGGA 706
QY 713 GAAAGTTATGATGAGCTGAAACATTTCTCTCAAAAAATATTTCAGGAGAGCGCTTGCAGAAA 772
DB 707 GAAATTTATGATGAGGCTGAAATTTCTCTCAAAAAATATTTCAGGAGAGCGCTTGCAGAAA 766
QY 773 GATTCCGCGCATCCAGTATATCTTTCACTAGAGATACCGGACGCTTGTGAAATATGTTGCGCA 832
DB 767 GATTCCGCTCTCCAGT---CTTTCCGCGAGAGATCGGGGACGTTTGGAAATATGTTGCGCA 823

```

QY 833 CACCAATTTGCCACGCTTGGRAAGCAAGAAATTACATGGACGCTCTTTGGACAGCACACTAA 892  
DB |||||  
834 CACATATTTGGCGCAITTTGGAGCAAGGAATTTACATCCAAAGTCTTTTGGACAGCACACTGA 883  
QY 893 AAATAAGAACGC-----CGCCGAGAAACCTTTTGAACCTTGCAAAATTGCAATTCAA 943  
DB |||||  
884 GAACACGAAGTCATATGTGAAGAGCAAAAACCTTTTGAACCTCGCAAAATTGGAGTTCAA 943  
QY 944 TATATTTCACTCTCTTACAGAGAGAGAGTTAAAAACATGTTTCCCGATGGTGGAAAGACTC 1003  
DB |||||  
944 CATCTTTCAATCTTTACAAAAGAGGAGTTGAAAAGTCTGGTCAGATGGTGGAAGAAATC 1003  
QY 1004 GGGTTCTCTGAGATGACCTTCTGTGACATCGTCACGTGGATGATACACTACGCTTTGGCTTC 1063  
DB |||||  
1004 GGGTTTTCTGAGATGACCTTCTGCGACATCGTCACGTGGAAATACACTTTTGGCTTC 1063  
QY 1064 CTGCATTTGGTTTCGAGCCTCAACATTTCTGGATTACAGCTCGGCTTTTACCAAGATGCTCA 1123  
DB |||||  
1064 CTGCATTTGGTTTCGAGCCTCAACATTTCTGGATTACAGCTCGGCTTTTACCAAGATGCTCA 1123  
QY 1124 TCTTATCAGGTTCTTGAAGACATGTACAGCTCTTTGGGACAGTACGCTGGAAT 1183  
DB |||||  
1124 TCTTATCAGGTTCTTGAAGACATGTACAGCTCTTTGGGACAGTACGCTGGAAT 1183  
QY 1184 CTTTACAGCGCAATTTAAGAGATGGATCCGTCGCGATGGAATGCTTCCAGATATAT 1243  
DB |||||  
1184 CTTTACAGCGCAATTTAAGAGATGGATCCGTCGCGATGGAATGCTTCCAGATATAT 1243  
QY 1244 GAAAGGAGTGATCATGATGGTTTATCACACCGCTAAATGAATGGCTCGAGTGGCAGAGAA 1303  
DB |||||  
1244 GAAAGGAGTGATCATGATGGTTTATCACACCGCTAAATGAATGGCTCGAGTGGCAGAGAA 1303  
QY 1304 GGCTCAAGCGCGAGACACGCTCAACTATGCAAGACAGCGCTTGGAGGCGTGTATTGATTC 1363  
DB |||||  
1304 GGCTCAAGCGCGAGACACGCTCAACTATGCAAGACAGCGCTTGGAGGCGTGTATTGATTC 1363  
QY 1364 GTATATGAGAGCAAGCAAGTGGATCGCCACCTGGTTATCTGCCAGTTTGAGGATGACTT 1423  
DB |||||  
1364 GTATATGAGAGCAAGCAAGTGGATCGCCACCTGGTTATCTGCCAGTTTGAGGATGACTT 1423  
QY 1424 GGAGAACGGGAAAGTTAGCTCTGTCTCATCGGCCATGCGCACTGCAACCCATTCTGACGTT 1483  
DB |||||  
1424 CGAGAAATGSGAAGATTAGCTGTGGTTCATCGCATATCCGCATTTGCACCCATTCTGACAA 1483  
QY 1484 GGACATCCCTTTCTGTATCATCTCTCAAGGAAGTTGACTTCCCATCGAAGCTCAATGA 1543  
DB |||||  
1484 GGACATCCCTTTCTGTATCATCTCTCAAGGAAGTTGACTTCCCATCGAAGCTTAACGA 1543  
QY 1544 CTTGATATGTATCATCTCTGTATTAAGAGTGATACAGGTGCTACAGGCGAGACAGGC 1603  
DB |||||  
1544 CTTGGCATGTCATCTCTGTATTAAGAGTGATACAGGTGCTACAGGCGAGACAGGC 1603  
QY 1604 CCGTGGAGAAAGCTTCTGTATCATCTGTATTAAGAGCAATCTCGATTAACGGA 1663  
DB |||||  
1604 TCGTGGAGAAAGCTTCTGTATCATCTGTATTAAGAGCAATCTCGATTAACGGA 1663  
QY 1664 AGAAGATGCTCTGATCATATCAACTTCATGATCAGGACCGCAATCAGAGAAATTAATTG 1723  
DB |||||  
1664 GGAAGATGCTCTGATCATATCAAGGCGCAATGATCAGTGACGCTCAAGAGTTAAATTG 1723  
QY 1724 GGAGCTTTAAAGCGAGACACAGTGTTCCTCATCTTCCATGCTTCCAGAAACACGCTTTGACAT 1783  
DB |||||  
1724 GGAATCTTCAACCGAGACATCAATGTTCCCATCTCGCGAAGAAACATGCTTTTGGACAT 1783  
QY 1784 AAGCAGAGTTTGGCATCAGGTTTACAGATACCGAGATGGCTTACAGCTTTTGCACAGTTGA 1843  
DB |||||  
1784 CGCCAGAGCTTTTCATTAAGGCTTCAAAATACCGAGAGGCTTACAGGCTTGCACAGTTGA 1843  
QY 1844 AACAAAGAGTTTGGTATGAGAACCGCTCATGTAACCTGTGCTTTGTAAACAACTTCAA 1903  
DB |||||  
1844 AACGAAGAGTTTGGTATGAGAACCGCTCATGTAACCTGTGCTTTGTAAACAACTTCAA 1903  
QY 1904 ATCTACAATATTAATCTGAGGATGCCCTATGGGTGTATATAGGCGACACAAAAATAATAT 1963

DB 1904 ATCTATGCCCTATGCTATGTCGGGTTAAAAATATATGTGGAAGCTAGCCGTTGGATGTAGA 1963  
QY 1964 GGTTCGTGTAGTAAGCTGTAAATTTATGAAAAAATAAAAAAATAAAAAA 2013  
DB |||||  
1964 GGAAGAAGTTGTTATTAATTAAGAGTTGTAATTTAAAAAATAAAAAAATAAAAAA 2013  
RESULT 6  
US-10-041-007-21  
; Sequence 21, Application US/10041007  
; Publication No. US20020164736A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P.T.  
; APPLICANT: Schepmann, Hala G  
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
; FILE REFERENCE: P02081US1  
; CURRENT APPLICATION NUMBER: US/10/041,007  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259,881  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21  
; LENGTH: 2018  
; TYPE: DNA  
; ORGANISM: Abies grandis  
US-10-041-007-21  
Query Match 64.9%; Score 1306.8; DB 14; Length 2018;  
Best Local Similarity 81.0%; Pred. No. 0;  
Matches 1595; Conservative 0; Mismatches 342; Indels 33; Gaps 5;  
QY 68 CAGGTTCGTGCTCAGTTCCTTCATGAGATTAAGCTCTCCGTAGAACAAATCCCAACTCT 127  
DB |||||  
53 CAATCGTGTGATGAGTCTACCCATGAGCTTAAGGCTCTCTCTAGAACAAATCCCAACTCT 112  
QY 128 TGAATCTGACGGCGGGGAAATCCGTCGCAATTCATAAACATGCTTTTGACAAGCGT 187  
DB |||||  
113 AGGAATGAGTAGGCGAGGAAATCTATCACTCTCTCCATCAGCATGAGCTTACCACCGT 172  
QY 188 CGCATCTACTGATCTGTACAGAGACGGTGGGCAATCATCATTTCCAACTGCGGAGCA 247  
DB |||||  
173 TGTAACCGATGATGCTGTACGAAGACGATGGGGATTTCCATTCACACTCTCGGAGCA 232  
QY 248 TGATTTATACAGTCTCTGATCTCAACGCCCTTATGGAGCACTGATACCGGGAACGTC 307  
DB |||||  
233 TGATGTATACAGTCTCT---TTACCAAGGCTTATGAGAAATACTGATACCTGGAGCGTGC 289  
QY 308 TGACAGACTTATTTGGGGAAGTAAAGGATATAAATGTTCAATTTCAAGTGGCTGGAGATGG 367  
DB |||||  
290 TGAGAAACTGATCGGGGAAGTAAAG---AACATGTTCAATTCGATGTCATTAGAAGATGG 346  
QY 368 AG-----GCAATGATCTCTCAACGACTTTTGGTGGTGGTGGTGGTGGTGGTGGTGG 412  
DB |||||  
347 AGGTTAATGAGTCCGCTCAATGATCTCAACGCTTTTGGATTTGTCGACGCTTGA 406  
QY 413 ACGTTTGGGATCGACAGGCAATTTCAAAAGAGATATAAAGCAACGCACTCGGATTTGTTAA 472  
DB |||||  
407 ACGTTTGGGATCCATAGACATTTCAAGATGAGATATAAATCGGCGCTTGATTTGTTTA 466  
QY 473 CAGTTATGGAACGAAAAAGGCAATTTGGATTTGGAGGGAGAGTGTGTGACTGACCTCAA 532  
DB |||||  
467 CAGTTATTTGGGCGAAAAATGGCATCGGATCGGGAGGAGAGTGTGTGTACTGATCTGAA 526  
QY 533 CTCAACCGCTTTGGGCTTCGAATCTCCGACTACAGGATACACTGTGTCTCAGATGT 592  
DB |||||  
527 CTCAACTGCGTGGGGCTTCGAACCTTACGACTACACGAGTACCGGCTGTCTTCAAGATGT 586  
QY 593 TTTCAACGCTTTTAAAGACAAAAATGGCAATTTTCTCCACTGCCAATATTCAGATAGA 652  
DB |||||  
587 TTTCAACGCTTTCAAGGCCAAATGGGCAAGTTTCTCTGCTCTGAAATATTCAGACAGA 646

QY 653 GGGAGAGATTAGAGCGTTCTCAATTTATTTCAGGCGCTCCCTCGTGCCTTTCCCGCGCA 712  
 Db 647 TGAAGAGATCAAGAGCGTTCTGAATTTATTCGGGCGCTCCCTCATTTGCTTTCCAGGGA 706  
 QY 713 GAAAGTTATGGATGAAGCTGAACCAFTCTCTACAAATATTTAAGAGAGACCGCTGCAAAA 772  
 Db 707 GAAATTTATGGATGAGCGTGAATCTCTCTACCAATATTTAAAGAGAGACCGCTGCAAAA 766  
 QY 773 GATTCGGGATTCAGTATATCTTCTACATAGATACGGGACGTTCTCGAATATGGTTGCA 832  
 Db 767 GATTCGGGATTCAGTATATCTTCTACATAGATACGGGACGTTCTCGAATATGGTTGCA 823  
 QY 833 CACCAATTTGCCACGCTTGAAGCAAGGAATACATGAGCGTCTTTGGACAGACACTAA 892  
 Db 824 CACATATTTGCCCGGATTTGAAGCAAGGAATACATCCAGCTCTTTGGACAGACACTGA 893  
 QY 893 AATAAGAAACGCG-----CGCCGAGAACTTTTAGAACTTTGCAAAATTTGAAATTCAA 943  
 Db 884 GAACACGAAGTCATATGTGAAGAGCAAAAACCTTTTAGAACTCGCAAAATTTGAGTTCAA 943  
 QY 944 TATATTTCACTCCTTACAGAGAGAGATTAAACATGTTTCCGATGTTGGAAAGACTC 1003  
 Db 944 CATCTTTCAATCCTTACAAAGAGAGGAGTTAGAAAGTCTGGTCAGATGTTGGAAAGACTC 1003  
 QY 1004 GGGTTCTCTGAGATCACCTTCTGTCGACATCTCGTGGAACTACTACGCTTTGGCTTC 1063  
 Db 1004 GGGTTTTCTGAGATGACCTTCTGCGGACATCTGTCAGTGGAACTACTACACTTTGGCTTC 1063  
 QY 1064 CTGCATTGCTGAGCGCTCAACATTTCTGATTCAGATCGGCTTTTACCAAGATGCTCA 1123  
 Db 1064 CTGCATTGCTGAGCGCTCAACATTTCTGATTCAGATCGGCTTTTACCAAGATGCTCA 1123  
 QY 1124 TCTTATCAGGTTCTTGAAGAGATGAGGAGTTAGAAAGTCTGGTCAGATGTTGGAAAGACT 1183  
 Db 1124 TCTTATCAGGTTCTTGAAGAGATGAGGAGTTAGAAAGTCTGGTCAGATGTTGGAAAGACT 1183  
 QY 1184 CTTTACAGGACAAATTAAGAGATGGGATCGTCCGCGATGGAATGCTTCCAGAAATAT 1243  
 Db 1184 CTTTACAGGACAAATTAAGAGATGGGATCGTCCGCGATGGAATGCTTCCAGAAATAT 1243  
 QY 1244 GAAAGAGGTATCATGATGTTTATCACCGGTAAATGGAATGCTTCCAGAGTGCAGAGAA 1303  
 Db 1244 GAAAGAGGTATCATGATGTTTATCACCGGTAAATGGAATGCTTCCAGAGTGCAGAGAA 1303  
 QY 1304 GCGTCAAGCGGAGACGCTCAACTATGCAAGACAGGCTTGGAGCGGTTTGTGATTC 1363  
 Db 1304 GCGTCAAGCGGAGACGCTCAACTATGCAAGACAGGCTTGGAGCGGTTTGTGATTC 1363  
 QY 1364 GTATATGCGAGAAAGTGGATCGCCACTGGTTTATCTGCCACGTTTGGAGGATCTT 1423  
 Db 1364 GTATATGCGAGAAAGTGGATCGCCACTGGTTTATCTGCCACGTTTGGAGGATCTT 1423  
 QY 1424 GGAAGCGGAAAGTTAGCTCTGCTATGCTCCCATGCGGACTGCAACCCATTTGAGCTT 1483  
 Db 1424 GGAAGCGGAAAGTTAGCTCTGCTATGCTCCCATGCGGACTGCAACCCATTTGAGCTT 1483  
 QY 1484 GGAATCCCTTTCTGATCATATCTTCAAGGAAGTTGACTTCCATCAAGCTTAACGA 1543  
 Db 1484 GGAATCCCTTTCTGATCATATCTTCAAGGAAGTTGACTTCCATCAAGCTTAACGA 1543  
 QY 1544 CTTGATATGATCATCTTCCGATTAAGAGTGTATACCGTGTCTTCAAGGACAGAGGC 1603  
 Db 1544 CTTGATATGATCATCTTCCGATTAAGAGTGTATACCGTGTCTTCAAGGACAGAGGC 1603  
 QY 1604 CCGTGGAGAGAGCTTTCGCTATATCATCTTATATGAAAGACAACTCTCGAATTAACGA 1663  
 Db 1604 CCGTGGAGAGAGCTTTCGCTATATCATCTTATATGAAAGACAACTCTCGAATTAACGA 1663  
 QY 1664 AGAAGATGCTTCTGAATCATATCACTTTCATGATCAGGAGCGCAATCAGAGAAATTAATG 1723  
 Db 1664 AGAAGATGCTTCTGAATCATATCACTTTCATGATCAGGAGCGCAATCAGAGAAATTAATG 1723  
 QY 1724 GGAGCTTCTTAAAGCCGACAAAGTGTTCCTCATCTTCCAGAAACACGCAATTTGACAT 1783

Db 1724 GGAATCTTCAAAACAGACATCAATGTTCCCATCTCGGGAAGAAACATGCTTTTGACAT 1783  
 QY 1784 AAGCAGAGTTTGGCATCAGGTTACAGATACCGAGATGGCTACAGCTTTTGCCAAACGTTGA 1843  
 Db 1784 CCGCAGAGCTTTCCATTCAGGCTACAAATACCGAGACGGCTACAGGTTTGCCAAACGTTGA 1843  
 QY 1844 AACAAGAGTTTGGTATGAGAACCGTCAATGAAACCTGTGCGCTTTTAAACAACACTTCAA 1903  
 Db 1844 AACAAGAGTTTGGTATGAGAACCGTCAATGAAACCTGTGCGCTTTTAAACAACACTTCAA 1903  
 QY 1904 ATCTACAATATTAATGAGAGATGCGCTATGCGGTGTATATAGGGCAACAAAAATAATAT 1963  
 Db 1904 ATCTATGCGCTATGCTATGTCGGGTTAAATATATGCGAAGGTAGCGTTGGATGTAGA 1963  
 QY 1964 GGTGTGTTAGTAAAGCTGTAATTTATGAAAAAATAAAAAAATAAAAAA 2013  
 Db 1964 GGAATAAGTTTGTATTAATTAATAAGTTGTAATTTTAAAAAATAAAAAA 2013  
 RESULT 7  
 US-10-025-145A-3  
 ; Sequence 3, Application US/10025145A  
 ; Publication No. US20030175861A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bohlmann, Rodney B.  
 ; APPLICANT: Steele, Christopher L.  
 ; APPLICANT: Phillips, Michael A.  
 ; TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)  
 ; FILE REFERENCE: WSUR118414  
 ; CURRENT APPLICATION NUMBER: US/10/025,145A  
 ; CURRENT FILING DATE: 2002-06-28  
 ; PRIOR APPLICATION NUMBER: US 09/360,545  
 ; PRIOR FILING DATE: 1999-07-26  
 ; PRIOR APPLICATION NUMBER: PCT/US98/14528  
 ; PRIOR FILING DATE: 1998-07-10  
 ; PRIOR APPLICATION NUMBER: US 60/052,249  
 ; PRIOR FILING DATE: 1997-07-11  
 ; NUMBER OF SEQ ID NOS: 107  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 3  
 ; LENGTH: 2018  
 ; TYPE: DNA  
 ; ORGANISM: Abies Grandis  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (6)..(1892)  
 ; OTHER INFORMATION:  
 US-10-025-145A-3  
 Query Match 64.9%; Score 1306.8; DB 15; Length 2018;  
 Best Local Similarity 81.0%; Pred. No. 0;  
 Matches 1595; Conservative 0; Mismatches 342; Indels 33; Gaps 5;  
 QY 68 CAGGTGCTGCTCAGTTCTTCTCATGAGATTAAGGCTTCTCGTAGAACAAATCCCAACTCT 127  
 Db 53 CAAATCGTTGATCAGTTCTTACCCATGAGCTTAAGGCTCTCTTAGAACAAATCCAGCTCT 112  
 QY 128 TGAATCTCAGGCGCGGGAATCCGTCGCGCATTCATAAACATGTTGTTGACAAGCGT 187  
 Db 113 AGGAATGAGTAGGAGGAGGAAATCTATCACTCTTCCATCAGCATGAGCTCTACCCGT 172  
 QY 188 CGCATCTACTGATTTCTGTACAGAGACGCGTGGGCAATCATTCACCTCTGGGACGA 247  
 Db 173 TGTAAACGATGATGGTGTACGAAGACGATGGCGGATTTCCATTCACCTCTGGGACGA 232  
 QY 248 TGATTTCAATCAGTCTCTGATCTCAAGCCTTATGGAGACCTGATTTACCGGAAACGTCG 307  
 Db 233 TGATGTCATACAGTCTCTCTTACCAACGCGCTTATGAGAAAAATCGTACCTGGAGCGTCG 289  
 QY 308 TGACAGACTTATGGGGAAGTAAAGGATATAATGTTCAATTTCAAGTCGCTGGAGATGG 367

290 TGAGAACTGATCGGGGAAGTAAAG---AACATGTTCAATTCGATGTCATTTAGAGATGG 346  
368 AG-----GCAATGATCTCTTCAACGACTTTTGTGCTGGTCAATGAGCGTTGA 412  
347 AGAGTAAATGATCGCGCTCAATGATCTCATTCACCGCTTTGGATTGTGCGACGCTTGA 406  
413 ACGTTTGGGAATCGACAGGCAATTTCAAAAAGAGATAAACGCGCACTCGATATGTTAA 472  
407 ACGTTTGGGGATCCATAGACATTTCAAGATGAGATAAATCGGCGCTTGATATGTTTA 466  
473 CAGTTATTTGGAACGAAAAAGCATTTGGATGTGGGAGGAGAGTGTGTGACTGACCTCAA 532  
467 CAGTTATTTGGGCGAATATGCAATCGGATCGGAGGGAGAGTGTGTACTGATCTGAA 526  
533 CTAAACCGCTTGGGCTTGGAACTCTCCGACTACACGGATACACTGTGTCTTCAGATGT 592  
527 CTAACTCGGTGGGCTTCAACCGCTTACGACTACACGGATACCGGCTGTCTTCAGATGT 586  
593 TTTGACGCTTTTAAAGACAAAAATGGCAATTTTCTCCACTGCCAATATTCAGATAGA 652  
587 TTTCAAGCTTTCAAGGCGCAAAATGGGCAATTTTCTCTCTGCTGCTGAAAAATATTCAGACAGA 646  
653 GGGAGAGATTAGAGCGCTTCTCAATTTATTCAGGCGCTCCCTCGTCGCTTTCCCGGGA 712  
647 TGAAGAGATCAGAGCGCTTCTGAATTTATTCGGGCTTCCCTCATTTGCTTTCCAGGGGA 706  
713 GAAAGTTATGATGAGCTGAAACATTTCTCAAAAATTTTAAAGAGAGCGCTTGCAAAA 772  
707 GAAAAATATGATGAGGCTGAAATCTTCTCTACCAAAATTTTAAAGAGAGCGCTTGCAAAA 766  
773 GATTCGGCATCCAGTATACATTTCTACTAGATAGAGGAGCTTCTGGAATATGTTGCGCA 832  
767 GATTCGGCTCTCCAGT-----CTTTGCGGAGATCGGGGACGTTTGGAAATATGTTGCGCA 823  
833 CACCAATTTGCGAGCTTGGAGCAAGGAATTTACATGGACGCTTTTGACAGCACACTAA 892  
824 CACATATTTGCGGATTTGGAAGCAAGGAATTTACATCAAGTCTTTGACAGGACACTGA 883  
893 AAATAGAACGCG-----CGCGGAGAACTTTTGAAGCTTTCGAAAAATTTGGAATTCAA 943  
884 GAACACGAAGTCATATGTGAAGAGCAAAAACTTTTGAAGCTCGCAAAATTTGGAGTTCAA 943  
944 TATATTTCACTCTTCAAGAGAGAGATTTAAACATGTTTCCGATGGTGGAAAGACTC 1003  
944 CATCTTTCAATCTTTACAAAAAGAGGAGTTAGAAAGTCTGTGTCAGATGGTGGAAAGAAATC 1003  
1004 GGGTTCTCTCAGATGACCTTCTGTGCAATCGTTCACGTGGAATACACTACGCTTTGGCTTC 1063  
1004 GGGTTTCTCAGATGACCTTCTGCGACATCGTTCACGTGGAATACACTACACTTTGGCTTC 1063  
1064 CTGCAATTCGCTGAGCCTCAACATTTCTGGATTGAGTCTCGGCTTTACCAAGATGTCTCA 1123  
1064 CTGCAATTCGCTGAGCCTCAACATTTCTGGATTGAGTCTCGGCTTTGCAAGAGCTGTCA 1123  
1124 TCTTATCAGCGTCTTTCAGACATGTACAGCTCTTGGACAGTACAGTCTGCGCTTTGCGAACT 1183  
1124 TCTTATCAGCGTCTTTCAGACATGTACAGCACTCTTGGACAGTACAGTCTGCGCTTTGCGAACT 1183  
1184 CTTTCAAGCGCAATTAAGAGATGGGATCCGCTCGCGATGGAATGCGCTTCCAGAAATATAT 1243  
1184 CTTTCAAGCGCAATTAAGAGATGGGATCCGCTCGCGATGGAATGCGCTTCCAGAAATATAT 1243  
1244 GAAAGGAGTGTACATGATGGTTTATCACCGGTAAATGAAATGGCTCGAGTGGCAGAGAA 1303  
1244 GAAAGGAGTGTACATGATGGTTTATCACGACCGGTAAATGAAATGGCTCGAGAGGCGAGGA 1303  
1304 GGCTCAGGCGGACAGACGCTCAACTATGCAAGACAGCGCTTGGAGGCGCTTTTGTGATTC 1363  
1304 GGCTCAGGCGGACAGACGCTCAACTATGCTCGGAGAGCTTGGAGGCGCTTTATTTGATTC 1363  
1364 GTATATGAGGAAGCAAGTGGATCGCCACTGGTTTATCTGCCACGCTTTGAGGAGTACTT 1423  
1364 GTATATGCAAGAGCAAGTGGATCGCCACTGGTTTATCTGCCACTGGTTTGTGATGAGTACTA 1423

QY 1424 GGAGAACGGGAAGTTAGTCTGTCTCAPGCCCTGCGCACTGCAACCCATCTTGACGTT 1483  
Db 1424 CGAATTCGGGAAGTTAGTCTGTGTGTCATCGCATATCCGCAATTCGCAACCCATCTTGACAA 1483  
QY 1484 GGACATCCCTTTCTCTGATCATCTCTCAAGGAAGTTCAGTCCCATCGAAGCTCAATGA 1543  
Db 1484 GGACATCCCTTTCTCTGATCATCTCTCAAGGAAGTTCAGTCCCATCGAAGCTTAAAGGA 1543  
QY 1544 CTTGATATGATCATCTCTTCGATTAAAGAGGTGATACCGGTGCTACAAGGCGAGCAGGCG 1603  
Db 1544 CTTGCGCATGCGCATCTCTTCGATTACGAGGTGATACCGGTGCTACAAGGCGAGCAGGCG 1603  
QY 1604 CCGTGGAGAGAAGCTTCTCTATATCATGTTATGAAAGACAAATCTCTGGATTAAACGGA 1663  
Db 1604 TCGTGGAGAAGAAGCTTCTCTATATCATGTTATGAAAGACAAATCTCTGGAGTATCAGA 1663  
QY 1664 AGAAGATGCTCTGATCATATCAACTTCATGATCAGGAGCGCAATCAGAGAAATTAATTTG 1723  
Db 1664 GGAAGATGCTCTCGATCATATCAACGCCATGATCAGTGACGTAATCAAGGATTAATTTG 1723  
QY 1724 GGAGCTTCTTAAAGCGAGACACAGTGTTCCTCATCTTCCCAAGAAAACACGCAATTTGACAT 1783  
Db 1724 GGAATCTTCAACGAGACATCAATGTTCCCATCTCGCGGAGAAACATGCTTTTGTGACAT 1783  
QY 1784 AAGCAGAGTTTGGCATCACGTTTACAGATACCGAGATGGCTTACAGCTTTTGCCAACTTGA 1843  
Db 1784 CGCCAGAGCTTTCATTTACGGCTACAAATACCGAGAGCGCTTACAGCGTTGCCAAGCTTGA 1843  
QY 1844 AACAAAGAGTTTGGTGTAGAGAACCGCTCATTTGAACCTGTGCTTTGTAACAACTTCAA 1903  
Db 1844 AACAAAGAGTTTGGTGTAGAGAACCGCTCATTTGAACCTGTGCTTTGTAACAACTTCAA 1903  
QY 1904 ATCTACAATATTAATGAGGATGCCCTATGCGTGTATATGGGCACACAAAAATAAATAT 1963  
Db 1904 ATCTATGCCCTATGCTATGTCGGGTTAAATATATGTGGAAGGTAGCGCTTGGATGTAGA 1963  
QY 1964 GGTTGTGTTAGTAAAGCTGTAAATTTATGAAAAAATAAATAAATAAATAAATAAATAA 2013  
Db 1964 GGATAAGTTTGTATTAATTTAATAAGTTGTAAATTTAATAAATAAATAAATAAATAAATAA 2013

## RESULT 8

US-10-025-145A-77  
; Sequence 77, Application US/10025145A  
; Publication No. US20030175861A1  
; GENERAL INFORMATION:  
; APPLICANT: Croteau, Rodney B.  
; APPLICANT: Bohmann, Joerg  
; APPLICANT: Steele, Christopher L.  
; APPLICANT: Phillips, Michael A.  
; TITLE OF INVENTION: Monoterpene Syntheses from Grand Fir (Abies Grandis)  
; FILE REFERENCE: WSUR118414  
; CURRENT APPLICATION NUMBER: US/10/025.145A  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: US 09/360,545  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: PCT/US98/14528  
; PRIOR FILING DATE: 1998-07-10  
; PRIOR APPLICATION NUMBER: US 60/052,249  
; PRIOR FILING DATE: 1997-07-11  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 77  
; LENGTH: 1890  
; TYPE: DNA  
; ORGANISM: Abies Grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1890)  
; OTHER INFORMATION:  
US-10-025-145A-77

Query Match		54.3%; Score 1092.4; DB 15; Length 1890;
Best Local Similarity		76.8%; P-adj. No. 8 6e-313;
Matches 1439; Conservative		0; Mismatches 396; Indels 39; Gaps 7;
QY	46	TTTCTATTACTCGCTGGTTTCCAGGTCGTCAGTCTCTCTCATGAGATTAAAGGCTC 105
DB	26	TGCTTTCCAAATCGGTCCTGCAAAATCGTGGATCGTTTCTACTTATGAGCATAAAGCTA 85
QY	106	TCGGTAGAATCCCAATCTTTGGAATCTGACGCGGGGGAATCCGTCGCGATTCCA 165
DB	86	TCAGTAGAATCCCAATCTTTGGAATCTGACGCGGGGGAATCTGTCACATCTCC 145
QY	166	TAAACATGTTTGAACAGCTCGCATCTACTG---ATTCTGTACAGAGACGCGTGGCA 222
DB	146	TGGAATGATTTGACCAACGAGTCTCTGATCATGTTGTACAAAGCGCATAGTCG 205
QY	223	ACTATCATTTCCAACTGTGGAGCATGATTTTATACAGTCTCTGATCTCAACGCCATTATG 282
DB	206	AGTTTCATTTCCAACTGTGGAGCATGATTTTATACAAATCTC---TATCAACGCCATTATG 262
QY	283	GAGCACTGATTTACCGGGAACGTCGACAGACTTATTTGGGAAGTAAAGGATATAATGT 342
DB	263	GGCACCTTCATACCGTGAACGTCGATAGACTTATTTGGGAAGTAAAGGATATA--T 319
QY	343	TCAATTTCAAGTCGCTGGAGATGAG---GCAATGATCTCTCTTCAAC 387
DB	320	TCACTTTCAATTTTCCAGCGAAGATGGAGAACTAATCACTCCCTCAATGATCTCATTTCAAC 379
QY	388	GACTTTTGTGTCGATGAGTTGAACGTTTGGGAATCGACAGGCAATTTCAAAAAGAGA 447
DB	380	GCCTTTTAAATGTCGATACGTTTGAACGTTTGGGAATTTGATAGACATTTCAAAAATGAGA 439
QY	448	TAAAGAGCACTCGATTTATGTAACAGTTTATTTGGAACGAAAGCATTTGGATGTTGGA 507
DB	440	TAAAGAGCACTAGACTATGTTTACAGTTTATTTGGAACGAAAGCATTTGGCAGTTGGA 499
QY	508	GGGAGAGTGTGTGACTGACCTCAACTCAACGCCCTTTGGGCTTTTGAACCTCTCCGACTAC 567
DB	500	GTGATAGTGTGTGCTGATCTCAACTCAACTGCCCTGGGTTTCGAAATTTCTTCGACTAC 559
QY	568	ACGATACACTGTGCTCTCAGATGTTTGAACGTTTAAAGCAAAATTTGGCA---AT 624
DB	560	ACGATACACTGTGCTCTCAGATGTTTGAACGTTTAAAGCAAAATTTGGCA---AT 619
QY	625	TTTCTCTCCACTGCGCAATATTCAGATAGAGGAGAGATTAGAGCGTTTCTCAATTTATTTCA 684
DB	620	AGTTGTATGTTTGGGCAATCCAAACAGAGAGAGATATAAGCGTTCTGAATTTATTTTC 679
QY	685	GGGCTCTCCCTGCTGCGCTTTCCGCGGAGAAATTTATGGATGAAGCTGAAACATTTCTCTA 744
DB	680	GGGCTCTCCCTCATTTGCTTTCTGCGGAGAAATTTATGGAGAGGCTGMAATCTTCTCTA 739
QY	745	CAAAATATTTAAGAGAGCCCTCAAAAGTTTCGGCATCCAGTATACTTTCTACTAGAGA 804
DB	740	AAATATATTATAAAGAGCCCTTCAAAATATTCTGTCTCCAGT---CTTTCACGAGAGA 796
QY	805	TACGGACGCTTCTGGAATATGTTTGGCACACCAATTTGCCACGCTTGGAGCAAGGAATT 864
DB	797	TAGATAGCTTCTGGAGATGTTTGGCAACAAATATGCCAATTTGGAACAGAGNACT 856
QY	865	ACATGACGCTTTTGGACAGCACACTAAAATTAAG-----AACGCCGCGAGAAAC 915
DB	857	ACATCATGATTATTTGGGAGAGAACGATCGTATGAGAGCTTATATATGAACATGGAGAAAC 916
QY	916	TTTTAGAACTTGCAAAATTTGGAATTTCAATATATTTTCACTCTCTTACAGAGAGAGATTAA 975
DB	917	TTTTAGAAATTTGCAAAATTTGGAATTTCAATATATTTTCACTCTCTTACAGAGAGAGCTAA 976
QY	976	AACATGTTTCCCGATGTTGGAAGACTTCGGGTTCTCTCTGAGATGACCTTCTGTCGACATC 1035
DB	977	AAGACCTCTCCAGATGTTGGAAGACTTCGGGTTCTCTCTGAGATGACCTTCTGTCGCGATC 1036
QY	1036	GTACAGTGGGAATACACTGCTTTGGCTTCTGTCATTTGGCTTCGAGCCTCAACATTTCTGGAT 1095

RESULT 9

US-10-025-145A-66  
; Sequence 66, Application US/10025145A  
; Publication No. US20030175861A1  
; GENERAL INFORMATION:  
; APPLICANT: Croteau, Rodney B.  
; APPLICANT: Bohlmann, Joerg  
; APPLICANT: Steele, Christopher L.  
; APPLICANT: Phillips, Michael A.  
; TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)  
; FILE REFERENCE: WSUR118414  
; CURRENT APPLICATION NUMBER: US/10/025,145A  
; CURRENT FILING DATE: 2002-06-28

DB	1037	GTCACTGGAAATTTACGCTCTGGCATCTTGCAATTTGAAACTGATCGCAAAACATTCGGAT 1096
QY	1096	TCAGACTCGGCTTTACCAAGATGTCTCATCTTATCAAGCTTCTTACGACATGTACGACG 1155
DB	1097	TCAGACTCGGCTTTGCCAAAATGTGTCTTATACCGGTTTGGACGATATATACGACA 1156
QY	1156	TCTTCGGCACAGTAGACGAGCTGGAATCTTTCACAGCAATTAAGAGATGGATCCGT 1215
DB	1157	CCTTTGGAACATGGAGGAGCTGGAATCTTCTACTGAGCATTTAAGAGATGGATCCGT 1216
QY	1216	CGCGATGGAATTCCTTCAGAAATATATGAAAGGAGTGTACATGATGTTTATCAACCG 1275
DB	1217	CTGCCACAGATTTGCTTCCAGAGTATATGAAAGGGTGTACATGGTGGTTTACGAACCG 1276
QY	1276	TAAATGAATGCTCGAGTGGCAGAGAGCTCAAGCCGAGACACGCTCAACTATGCA 1335
DB	1277	TAAATGAATGCTCGAGAGGAGCAAGTCTCAAGCCGAGAGAGCTCAACATGCTC 1336
QY	1336	GACAGCTTTGGGAGGCTGTTTGTATTCGTATATGAGGAAGCAAAAGTGGATCGCCACTG 1395
DB	1337	GACGAGCTTTGGGAGGCTTATCTTGTATGATGAAAGAGCTGATGGATCTCCAGTG 1396
QY	1396	GTTATCTGCCAGCTTTGAGGAGTACTTGGAGAACGGGAAAGTGTAGTCTGCTCATCGCC 1455
DB	1397	GTTATCTGCCAAGCTTTGAGGAGTACTGAGAACCCAGCAAGTGTGTTTGGTATTCGA 1456
QY	1456	CATCGCACTGCAACCCATTTCTGAGTGGACATCCCTTTCTGTATCATCTCTCAAGG 1515
DB	1457	TATTCGATTTGCAACCCATCTCTCATGATGATTTCCCTTACTCACCACTCTCGAGG 1516
QY	1516	AACTTCACTTCCCATCGAAGCTCAATGACTGTATGTATCATCTTCCGATTAAGAGTG 1575
DB	1517	AAATAGACTTTCCATTTGAGGTTTATGACTTAAATGTTTCCATCTCTCGACTTAAAGATG 1576
QY	1576	ATACAGCTGCTTAAAGGAGAGCAGGCGCGCTGGAGAAAGCTTCTGTATATCATGTT 1635
DB	1577	ACACTCGCTCTCAAGGCGGAGGCGCGCTGGAGAAAGCTTCTGTATATCATGTT 1636
QY	1636	ATATGAAGACACTCTGATTAACGGAAGAGATCTCTGAAATCATATCACTTCAATGA 1695
DB	1637	ATATGAAGAGAACTCTGATCAACAGAGAGAGATGCTATCAATCATATCAACGCTATGG 1696
QY	1696	TCAGGACGCAATCAGAGAAATTAATTTGGAGCTTCTTAAAGCCAGACAAACAGTGTCCCA 1755
DB	1697	TCAATTAATTAATCAAGAGAGTGAATTTGGAGCTTCTCCGACAGGACGGCACCGCTCATA 1756
QY	1756	TCATTTCCAGAAACACGCAATTTGACATAGAGAGATTTGGCATCAGGTTACAGATACC 1815
DB	1757	TTGCTTTCAGAAACACGCTTTTGACATCTTCAAGGTTTCCCTTTCAGCGCTACAAATACC 1816
QY	1816	GAGATGCTACAGCTTTGCCAAGCTTTGAAACAAAGAGTTTGGTGTAGAGAACCGCTCATTTG 1875
DB	1817	GAGATGGTTTACGCTTGGCTTCCCAACAGAAACCAAGAAATTTGGTGTAGAGAACAGTCTTGT 1876
QY	1876	AACCTGTGCTTTG 1889
DB	1877	AGTCTGTGCTTTG 1890

```
; PRIOR APPLICATION NUMBER: US 09/360,545
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: PCT/US98/14528
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: US 60/052,249
; PRIOR FILING DATE: 1997-07-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
; LENGTH: 2186
; TYPE: DNA
; ORGANISM: Abies Grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (34)..(1923)
; OTHER INFORMATION:
US-10-025-145A-66

Query Match      53.3%; Score 1072.4; DB 15; Length 2186;
Best Local Similarity 76.4%; Pred. No. 8.2e-307;
Matches 1428; Conservative 0; Mismatches 391; Indels 51; Gaps 7;

QY 83 TCTTCTCATGAGATTAAGGCTCTCCGTAGAACAAATCCCACTCTTGGAAATCTGCAGGCC 142
DB 90 TACTCATCATGAGCTCAAGCCTCTCGGCAGAACCAATCCCACTCTTGGAAATGTGTAGGG 149
QY 143 GGGGAAATCCGTGCGCATATCCATAAAACATGTGTTTGAACAGCGTCGATCTACTGATTC 202
DB 150 AGGGAAATCTTTACACACCTCTGTGAGCATGAGTTTGACCACCGCTGTATCTGATGATGG 209
QY 203 TGTACAGAGACGCGTGGCACTATCATCTCCAACTGTGGACGATGATTTTCATACAGTC 262
DB 210 TCTACAAAGACGATAGGTGACTATCAATCCAACTCTTGGACGACGATTTTCATACAGTC 269
QY 263 TCTGATCTCAACGCCCTTATGAGACACCTGATACCGGAAACGTGTGACAGACTTATTTGG 322
DB 270 TC---TATCAAGCCTTATGGGAGCCTTCTTACCGAGAACGTGCTGAGAACTGATTTGG 326
QY 323 GGAAGTAAAGATATATGTTTCAATTTCAAGTCGCTGGAAGATGAG-- 369
DB 327 GGAAGTGAAGGA--GATGTTTCAATTTCAATGCAATCGCAATCGGAAGATGAGAAATCAATGAGTCC 383
QY 370 --GCAATGATCTCTTCAACGACTTTTGTCTGTCGATGACGCTTGAACGTTTGGGAATCGA 427
DB 384 CCTCAATGATCTTATGAAAGACTTTGGATGTCATAGCGTTGAACGTTTGGGAGTTGA 443
QY 428 CAGGATTTCAAAAAGAGATAAAACGGCACTCGATTATGTTTAAACAGTTATTTGGAACGA 487
DB 444 TAGACATTTCAAAAAGAGATAAAATCAGCCCTTGATTATGTTTACAGTTATTTGGAACGA 503
QY 488 AAAAGGCAATTGATGTGGAGGGAGAGTGTGTGACTGACCTCACTCAACCGCCTTGGG 547
DB 504 AAAAGGTAATTGGATGCGGTAGAGATAGTGTGTTTCTGATGTCAACTGACCTGCTCGGG 563
QY 548 GCTTCGAACTCTCCGACTACACGGATACACTGTGCTTTCAGATGTTTGAACGTTTAA 607
DB 564 GTTTCGAACTCTTCCGCTACACGGATACAGTGTCTTTCAGAGGTTTGAAGTATTTCA 623
QY 608 AGACAAAATGGCAATTTTCTCCACTGCCAATATTCAGATAGAGGAGAGATTTAGAGG 667
DB 624 AGACAAAATGGGAGTTTTCATTTCTCTCC--TAGTACAAAAGAGAGAGACATCAGAAC 680
QY 668 GGTTCGAAATTTATTCAGGCGCTCCTCGTCCCTTTCGGCGGAGAAAGTTTGGATGA 727
DB 681 CGTTCGAAATTTATATCGGGGCTTCTTTCAITGCGCTTTCTCGGGGAGAAAGTTTATGAAGA 740
QY 728 AGCTGAAACATTTCTCAAAAATATTTTAAAGAGAGCCCTGCAAAAGATTTCCGGCATCCAG 787
DB 741 GGCTGAAATTTTCTTCAAGATATTTGAAAGAGCCGCTGCAAAAGATTTCCGGTCTCCAG 800
QY 788 TATACTTTCTACAGATACGGGAGCTTCTGGAATATGTTTGGCACACCAATTTGCCACG 847
DB 801 T---CTTTCACAGAATAGACTACACTTTTGAATATGTTTGGCACACAATATGCCAAG 857
```

```
QY 848 CTTTGAAGCAAGGAATTAATGAGCGTCTTTTGGACAGACACTA-----AA 893
DB 858 ATTGGAACAAGGAATTAATGAGTGTATTTGGACATCTTACAGTCCATGCTCAAGAA 917
QY 894 AATAAGAACGC-----CGCGAGAAACTTTTGAACCTTGCAAAATTTGGAATTTCAA 943
DB 918 GAAAAGGACGCAATATCTGGACAGCGAAAGCTTTTAGAATCTCGCAAAATTTGAGTTCAA 977
QY 944 TATATTTCACTCTTACAAAGAGAGAGTTTAAACATGTTTCCCGATGTTGTAAGAGCTC 1003
DB 978 CATCTTTCACTCTTCAACAGAGAGATTACAGTATCTCTCCAGATGGTGAATCATTC 1037
QY 1004 GGGTTCTCTGAGATGACCTTTCTGTGACATCGTCACTGGAATTAATACTACGCTTTGGCTTC 1063
DB 1038 GGGTTTGCCTGAACCTGACCTTTGGTTCGTCATCTGCACTGGAATTAATACTACGCTTC 1097
QY 1064 CTGCATTTGGTTCGAGCCTCAACATTTCTGGATTCTAGACTCGGCTTTACAGATGCTCA 1123
DB 1098 TTGCATTTGCGATGAGCCCAACATTTCTGCAATTCAGATTGGGCTTTGCCAAACGTTCA 1157
QY 1124 TCTTATCACTGTTCTTGACGACATGTAGACGTCTTTCGSCACAGTAGAGCTCGAAT 1183
DB 1158 TCTTATCACTGTTCTGGACGATCTCTAGACACTTTTCGAAAGATGATGAATCGAAT 1217
QY 1184 CTTCAAGCGCAATTAAGAGATGGATCCGTTCGCGATGGAATGCTTTCCAGAAATATAT 1243
DB 1218 CTTCAAGCGGAGCTTAGGAGATGGAATCCGTTCGAGAGAAAGACGCTCCAGAAATATAT 1277
QY 1244 GAAAGGATGTACATGATGTTTATCAACCGTAAATGAATGGCTCGAGTGGCAGAGAA 1303
DB 1278 GAAAGAAATCTACATGGCACTCTAGAAAGCCTTAACATGATGCGCGAGAGGCGAGAA 1337
QY 1304 GGCTCAAGCGCGAGACAGCTCACTATGCAAGACAGGCTTTGGAGGCGTGTGTTGATTC 1363
DB 1338 GACACAGGCGCGAGACAGCTCAATATGCTTGAAGGCTTGGAGAGTTATTCITGATTC 1397
QY 1364 GTATATGAGGAAGCAAGTGGATGCGCACTGGTTATCTGCCACGTTTGGAGGACTCTT 1423
DB 1398 GTATACACAAGAAAGTGGATGCGCAGCGGTTATCTGCCAACTTTTCGAGGAGTACTT 1457
QY 1424 GGAAACGGGAAAGTGTAGCTCTGCTCATCGCCCATGCGCACTGCAACCCATTTGACGTT 1483
DB 1458 AGAAGACGGAAGGTTAGCTCTGTCATCTGTCATGTCGAGCGGCATTGACACCCCTCTGACAT 1517
QY 1484 GGACATCCCTTCTCTGATCACAATCTCAAGGAAAGTTCGACTTCCATCGAAGCTCAATGA 1543
DB 1518 GGACGTACCGCTTCTCTGATGAGCTCTTGAAGGGAATAGATTTTCCATCGAGATTTATGA 1577
QY 1544 CTTGATATGATCATCTCTTCGATTAAGAGGTGATACAGGTGCTTACAGGCGAGACAGGC 1603
DB 1578 TTTGGCATCTTCTCTTACAGTAAAGAGGTGACACAGTGTACAGGCGAGACAGGGA 1637
QY 1604 CCGTGGAGAAAGCTTCGTCTATCATGTTTATGAAAGACAACTCTGATTAACCGGA 1663
DB 1638 CCGAGGAGAAAGCGTCAAGCATATCGTGTATGAAAGACAAATCCCGATTTAACAGA 1697
QY 1664 AGAAGATGCTCTGAATCATATCAACTTCATGATCAGGAGCGCAATCAGAGAAATTAATTTG 1723
DB 1698 GGAAGATGCTCTCAATCATATCAATGCAATGATCAACAGCATATCAAGAAATTAATTTG 1757
QY 1724 GGAGCTTCTTAAAGCGACAACTGTTCCCATCACTTCCAAAGAAACACGCAATTTGACAT 1783
DB 1758 GGAACTTCTCAAAACCGATAGCAATATTTCCAATGACTGCAACGGAACATGCTTATGAGAT 1817
QY 1784 AAGCAGAGTTTGGCATCAGGTTTACAGATCCGAGATGGCTTACAGCTTTCGCCACGTTGA 1843
DB 1818 AACCAGAGCTTTCACCAACTTTTACAAATATAGAGATGGCTTTCAGCGTTGCCACTCAAGA 1877
QY 1844 AACAAAGAGTTTGGTGTATGAGAACCGCTCATTTGAACCTGTGCTTTGTAACCAACTTTAA 1903
DB 1878 AACGAAAGTTTGGTGGAGGAGACGGTCTTGAACCAAGTCCCTTTTAACTTTAAACC 1937
```



```
Db 1664 TCGTGGTGAAGAGCTTCGTGTATATCATCTGTATATATGATGAAAGACAATCTCGGTATCAACCGA 1723
Qy 1664 AGAAGATGCTTGAATCATATCAACTTCATGATCAGGAGCGCAATCAGAGAATTAATTAATG 1723
Db 1724 AGAAGATGCCCTCAATCATATCAATGCCATGTCAATGACATAATCAAGAATTAATTAATG 1783
Qy 1724 GGAGCTTCTAAAGCCAGACCAAGTGTCCCATCATCTCCCAAGAAACAGCATTTGACAT 1783
Db 1784 GGAATCTTAAAGATCCAAAGCAATATTTCCAAATGCTGGCCAAAGAAACATGCTTTTGACAT 1843
Qy 1784 AAGCAGAGTGTGGCATCAACGGTTACAGATACCGAGATGCTACAGCTTTGCCAAAGTTGA 1843
Db 1844 AACAGAGCTCTCCACCATCTCTACATATATCGAGATGCTTTAGTGTTCACCAAGGA 1903
Qy 1844 AACAAAGAGTGTGGTGAAGAGAACCGTCATTGAACCTGTGCCCTTTGTAACACACTTCAA 1903
Db 1904 AACAAAAAAATGGTTATGGAACACACTCTTGAAATCTATGCTTTTAACTATAACCAT 1963
Qy 1904 ATCTACATATTAACCTGAGGATGCCCTATGGGTGTATATAGGACA 1950
Db 1964 TCCATAATAATAAGCTCATATGCTAAATATTATGGCCTTATGACATA 2010

RESULT 11
US-09-903-012-29
; Sequence 29, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1999-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 2196
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (69)...(1949)
; OTHER INFORMATION: myrcene synthase
US-09-903-012-29

Query Match 53.2%; Score 1071.6; DB 9; Length 2196;
Best Local Similarity 75.0%; Pred. No. 1.4e-306;
Matches 1430; Conservative 0; Mismatches 441; Indels 36; Gaps 6;

Qy 68 CAGTGTGCTGCTCAGTTCCTTCATGAGATTAAAGCTCTCCGTAGAACAAATCCCAACTCT 127
Db 116 CAAGTGTGATCAGTTCAATTCATGAAACATAGCCCTCCCTATAGAACAAATCCCAAAATCT 175
Qy 128 TGAATCTGACGGCGGGGAAATCCGTGCGCATTTCCAATAACATGTGTTTGACAAGCGT 187
Db 176 TGAATGCGTAGGCGAGGAAATCTGTACCGCTTCATGAGCATCAGTTTGGCCACCGC 235
Qy 188 CGCATCTACTGATCTGTACAGAGACGGGTGGCAACTATCATTTCCAACTGTGGGACGA 247
Db 236 TGCACCTGATGTGTGTACAAAGACGATAGGTGACTTACCATTCCAATATCTGGGACGA 295
Qy 248 TGAATTCATACAGTCTCTGATCTCAACGCCCTTATGGAGCACTGATTACCGGGAACTGTC 307
```

```
Db 296 TGATTTTCATACAGTCTC---TATCAACGCCTTATGGGAACCCCTCTTACCAGGAACGTGC 352
Qy 308 TGACAGACTTATTTGGGAGTAAAGGATATATGTTTCAATTTCAAGTCTGCTGGAAGATGG 367
Db 353 TGAGAGATTAAATTTGGGAGGTAAAGAAGATA---TTCAATTTCAATGTATGCTCGATGATGG 409
Qy 368 AGG-----CAATGATCTCTCTTCAACGACTTTTGTGGTCTGATGACGTGGA 412
Db 410 AAGATTAAATGAGTTCCTTTTAAATGATCTCATGCAACGCCTTTGGATAGTCTGATGAGTGA 469
Qy 413 AGTTTGGGAATCGACAGGCAATTTCAAAAAAGAGATAAAAAACGCACTCGATATATGTTAA 472
Db 470 ACGTTTGGGGATAGTAGACATTTCAAGAACAGATAAATCAGCTCTGGATTATGTTT 529
Qy 473 CAGTTTATTTGGAACAAAAGGCAATTTGGATGTGGGAGGAGAGTGTGTGACATGACCTCAA 532
Db 530 CCGTTACTGGGAGGAAACGGCAITGGATGTGGGAGAGACAGTATTTGTACTGATCTCAA 589
Qy 533 CTCAACCGCTTTGGGGCTTCGAACTCTCCGACTCACGGAATACACTGTCTCTTCAGATGT 592
Db 590 CTCAACTCGCTTTGGGGTTTCGAACTCTTCGATTACACGGGTACACTGTATCTCCAGAGT 649
Qy 593 TTTGNAACGTTTAAAGACAAAAATGGCAATTTTCTCCACTGCCAATATATTCAGATAGA 652
Db 650 TTTAAAGCTTTTCAAGATCAAAATGGACAGTTTGTATGCTCCTCCCC---GGTCAACAGA 706
Qy 653 GGGAGAGATTAGAGCGTCTCAATTTATTCAGGSCCTCCCTCGTCGCTTTCCTCCGGGA 712
Db 707 GGGTGAATCAGAGCGTCTTAACTTATTCGGGCTTCCCTCATGCTTCCCTCGGTGA 766
Qy 713 GAAAGTTATGATGAAGCTGAAACATTTCTCAAAAATTTTAAAGAGAGCCCTCAGAAA 772
Db 767 GAAAGTTATGGAAGAAGCTGAAATCTTCTCCACAAGATATTTGAAAGAAGCTCTACAAA 826
Qy 773 GATTCGGCATCCAGTATATCTTTCATGAGATACGGACGTTCTTGGAAATATGTTGGCA 832
Db 827 GATTCAGTCTCCGCT---CTTTCACAGAGATAAAGTTTGTATGGAATATGGCTGCA 883
Qy 833 CACCAATTTGCGACGCTTTGGAAGCAAGAAATTCATGAGACCTCTTTGGACAGCACATA- 891
Db 884 CACAAATTTGCCAAGATTGGGAAGCAAGAAATTCATGAGACACACTTGAAGAAGACACAG 943
Qy 892 -----AAAATAAGAACGCCGCGAGAAAATTTTAGAACTTGCAAAATTTGGAATTCAA 943
Db 944 TGCATGCTCAATAAAATGCTGGGAAGAGCTTTTAGAACTTGCAAAATTTGGATTCAA 1003
Qy 944 TATATTTCACTCTTACAGAGAGAGAGTTAAACATGTTTCCCGATGGTGGAAAGATC 1003
Db 1004 TATATTTAACTCTTTACAAACAAAAGGAATTAACAATATCTTTGAGATGGTGGAAAGATC 1063
Qy 1004 GGGTTCCTCGATGACCTTCTGTCGACATCGTCACGTGGAATACACTACGCTTTGGCTTC 1063
Db 1064 GGAATTCGCTAAATTTGACATTTGCTCGGCATCGTCATGTGGAATTTACACTTTGGCCTC 1123
Qy 1064 CTGCAATTTGCGTTTCGAGCCTCAACATTTGGAATTCAGACTCGGCTTTTACCAGATGTCTCA 1123
Db 1124 TTGTAATTTGCCATTTGACCCAAAACATTTCTGCAATTCAGACTAGGCTTCGCAAAATGTGTCA 1183
Qy 1124 TCTTATCAGCGTTCCTTACGACATGTACGAGCTTCTCGGACAGTACGAGCTGGAAT 1183
Db 1184 TCTTGTCACTAGTTTGTGACGATATTTTACGACACTTTTGGAAACGATTTGACGAGCTTGAAT 1243
Qy 1184 CTTTCAAGCAGCAATTAAGAGATGGGATCCGTCCGATGGAATGCTTCCAGAAATATAT 1243
Db 1244 CTTTCACTCTGCAATTAAGAGATGGAATTCATCAGAGATAGAACACCTTCCAGAAATATAT 1303
Qy 1244 GAAAGAGTGTACATGATGTTTATCACCCGTAATGAAATGAAATGAGTCTGAGTGGCAGAGAA 1303
Db 1304 GAAATGTGTACATGCTGCTGTTTGAACCTGTAAATGAACTGACACGAGAGGCGGAGAA 1363
Qy 1304 GGCCTAAGGCGGAGACACGCTCAACTATGCAAGACAGCTTTGGAGCGGTGTTTGTATTC 1363
```

Db 1364 GACTCAAGGAGAGAAACACTCTCAACTATGTTGGAAGGCTTGGAGGCTTATTTGATTC 1423  
 Qy 1364 GTATATCGAGGAGAGAAAGTGGATCCCACTGTTATCTGCCACGCTTTGAGGAGTACTT 1423  
 Db 1424 ATATATGGAAGAGAGCAAAATGATCTCTAATGGTTATCTGCCCAATGTTTGAAGAGTACCA 1483  
 Qy 1424 GGAGAACGGGAAAGTTAGCTCTGCTCATCGCCCATCGGCACATGCAACCCCATCTGACGTT 1483  
 Db 1484 TGGAATGGAAGAGTGGATCTGCTATCGGTAGCAACATTCGCAACCCCATCTGACTTT 1543  
 Qy 1484 GGACATCCCTTTCCTGATCAACATCCTCAAGGAGTTGACTTCCCATCGAAGCTCAATGA 1543  
 Db 1544 GAATGATGATCATCTCTCGATTAGAGGTGATACACCGTGCTACAAGGCGCATAGGGA 1603  
 Qy 1544 CTTGATATGATCATCTCTCGATTAGAGGTGATACACCGTGCTACAAGGCGCATAGGGA 1603  
 Db 1604 TTTGGCATCGTCTCTCTTGGCTAGAGGTGATACACCGTGCTACAAGGCGCATAGGGA 1663  
 Qy 1604 CGGTGAGAGAGAGCTTTCGTCTATATCATGTTATATGAAGACAAATCCTGGATTAACGGA 1663  
 Db 1664 TCGTGGTGAAGAGCTTTCGTCTATATCATGTTATATGAAGACAAATCCTGGATTAACGGA 1723  
 Qy 1664 AGAAGATGCTGATCATATCAATCAATGATGATGATGATGATGATGATGATGATGATGATG 1723  
 Db 1724 AGAAGATGCTGATCATATCAATCAATGATGATGATGATGATGATGATGATGATGATGATG 1783  
 Qy 1724 GAGCTTCTAAAGCCAGAGCAACAGTGTTCCTCATCTTCCCAAGAAACAGCAATTTGACAT 1783  
 Db 1784 GGAATCTTAAAGCCAGAGCAACAGTGTTCCTCATCTTCCCAAGAAACAGCAATTTGACAT 1843  
 Qy 1784 AAGCAGAGTTTGGCATCAAGTTTACAGATACAGATACAGATACAGATACAGATACAGATAC 1843  
 Db 1844 AACAGAGCTCTCCACCACTCTACATATATACATATACATATACATATACATATACATAT 1903  
 Qy 1844 AACAGAGTTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1903  
 Db 1904 AACAAAAAATTTGGTTATGGAACACTCTCTTGAATCTATGCTTTTTTAACTATAACCAT 1963  
 Qy 1904 ATCTAATATTAATTAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1950  
 Db 1964 TCCATAATAAAGCTCAATGCTAAATTAATTTGGGCTTATGACATA 2010

RESULT 12

US-09-900-797-29  
 ; Sequence 29, Application US/09900797  
 ; Publication No. US20030087406A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chappell, Joseph  
 ; APPLICANT: No. US20030087406A1, Joseph P.  
 ; APPLICANT: Starks, Courtney M.  
 ; APPLICANT: Manna, Kathleen R.  
 ; TITLE OF INVENTION: SYNTHASES  
 ; FILE REFERENCE: 07678-025001  
 ; CURRENT APPLICATION NUMBER: US/09/900,797  
 ; CURRENT FILING DATE: 2001-07-06  
 ; PRIOR APPLICATION NUMBER: US/09/398,395  
 ; PRIOR FILING DATE: 1999-09-17  
 ; PRIOR APPLICATION NUMBER: 60/130,628  
 ; PRIOR FILING DATE: 1999-04-22  
 ; PRIOR APPLICATION NUMBER: 60/150,262  
 ; PRIOR FILING DATE: 1999-08-23  
 ; NUMBER OF SEQ ID NOS: 58  
 ; SOFTWARE: Fast-Seq for Windows Version 3.0  
 ; SEQ ID NO 29  
 ; LENGTH: 2196  
 ; TYPE: DNA  
 ; ORGANISM: Abies grandis  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (69)...(1949)  
 ; OTHER INFORMATION: myrcene synthase  
 ; US-09-900-797-29

Query Match 53.2%; Score 1071.6; DB 10; Length 2196;  
 Best Local Similarity 75.0%; Pred. No. 1.4e-306;  
 Matches 1430; Conservative 0; Mismatches 441; Indels 36; Gaps 6;  
 Qy 68 CAGGTGTCCTCAGTTCCTCATGAGATTAAGGCTCTCCGTAGAACAAATCCCACTCT 127  
 Db 116 CAAAGTCGTTGATCAGTTCAATTCATGAACATAGCCCTCCCTATAGAACAAATCCCAATCT 175  
 Qy 128 TGGAAATCTGACGCCGGGGAATCCGTCCGGCATTCATAAATCATCTGTTTGAACAAGCT 187  
 Db 176 TGAATGCTGAGCGGAGGGAATCTGTACGCCCTTCCATGAGCATCAGTTTGGCCACCGC 235  
 Qy 188 CGCATCTACTGTTCTGTACAGAGACGCGTGGGCAACTATCAATTCACAACTGTGGACGA 247  
 Db 236 TGCACCTGATGTTGTGTACAAAGACGATAGTGACTACCAATTCACAAATCTGGACGA 295  
 Qy 248 TGAATTCATACAGTCTCTGATCTCAAGGCTTATGAGACACCTGATTAACCGGAAACGTGC 307  
 Db 296 TGAATTCATACAGTCTCTCTCAAGGCTTATGAGGAAACCTCTTACAGGAACGTGC 352  
 Qy 308 TGACAGACTTATTTGGGGAAGTAAAGGATATAATGTTTCAATTTCAAGTCGCTGGAAGATGG 367  
 Db 353 TGAGAGATTAATTTGGAGGTAAGAAGATA---TTCAATTCATGTACTCTGATGATGG 409  
 Qy 368 AGG-----CAATGATCTCTTCAACGACTTTTGTCTGGTGCATGACGTTGA 412  
 Db 410 AAGATTAAATGAGTTCCCTTTAAATGATCTCATGCAACGCCCTTGGATAGTCGATAGCGTTGA 469  
 Qy 413 ACGTTTGGGAATCGACAGGCAATTTCAAAAAGAGATAAAAACGGCACTGATTAATGTTAA 472  
 Db 470 ACGTTTGGGATAGCTAGACATTTCAAGAACGAGATAACATCAAGCTCTGATTAATGTTT 529  
 Qy 473 CAGTTATGGAACGAAAAAGGCAATTCGATGTGGAGGAGAGAGTGTGTGACTCACCTCAA 532  
 Db 530 CCGTTACTGGAGGAAAAACGGCAATGATGATGGAGAGACAGTATTTGTTACTGATCTCAA 589  
 Qy 533 CTCAACCGCTTGGGGCTTCGAACTCTCCGACTTACACGGATACACTGTGTCTTCAAGTGT 592  
 Db 590 CTCACCTGCTTGGGGTTCGAACTCTCGATACACGGGTACACTGTATCTCCAGAGGT 649  
 Qy 593 TTTGAAACGTTTTTAAAGACAAAAATGGGCAATTTTCTCCACTGCCAATATTCAGATAGA 652  
 Db 650 TTTAAAAAGCTTTTCAAGATCAAAATGGACAGATTTGATGCTCCCCC---GGTCAGACAGA 706  
 Qy 653 GGGAGAGATTAGAGGCGTTCTCAATTTTATTCAGGGCTCTCCGCTCGCCCTTTCCCGCGCA 712  
 Db 707 GGGTGAATCAGAGGCTTCTTAATTTATATCGGGCTTCCCTCATTTGCTTCCCTGGTGA 766  
 Qy 713 GAAAGTTATGGATGAAGCTGAAACATTTCTCTACAAAATATTTAAGAGAGAGCCCTGCAAAA 772  
 Db 767 GAAAGTTATGGAAGAAGCTGAAATCTTCTCCCAAGATATTTGAAAAGAAGCTCTACAAAA 826  
 Qy 773 GATTCGGCATCCAGTATATCTTCACTAGAGATACGGGACGTTCTGCAATATGTTGGCA 832  
 Db 827 GATTCAGTCTCCGCT---CTTTCAAGAGATAAAGTTTGTATGGAATATGCTGCGCA 883  
 Qy 833 CACCAATTTGCCACGCTTGGAGCAAGGAATATACATGGACGCTTTTGGACAGACACTA- 891  
 Db 884 CACAAATTTGCCAAGATTTGGAAGCAAGAAATATACATAGACACACTTTGAGAAAGACAC 943  
 Qy 892 -----AAATAGAAACGCCCGGAGAACTTTTAGAATCTTGCAAAATGGAATTCAA 943  
 Db 944 TGCATGGCTCAATAAAAAATGCTGGGAAGAGCTTTTAGAACTTTGCAAAATGAGTTCAA 1003  
 Qy 944 TATATTTCACTCTCTTCAAGAGAGAGAGTTTAAACATGTTTCCCGATGGTGGAAAGACTC 1003  
 Db 1004 TATATTTCACTCTCTTCAACAAAAGAAATTTACATATCTTTTGAAGTGTGGAAAGATC 1063  
 Qy 1004 GGGTCTCTGAGATGACCTTCTGTGACATCGTCAAGTGGAAATATCTACGCTTTGGCTTC 1063  
 Db 1064 GGAATTTGCCTAAATTTGACATTTGCTCGGCATCGTCTGGAATTTCTACACTTTGGCCTC 1123

```
QY 1064 CTGATTGGTTCGAGCTCAACATCTTGATTCAGATCGGCTTTACCAAGATGTCTCA 1123
Db 1124 TTGATTGCCATTGACCCAAACATCTCGATTCAGACTAGCTTCGCCAAATGTGCA 1183
QY 1124 TCTTATCAGGTTCTTGACGACATCTAGCAGCTCTTCGGCACAGTAGAGCTGGAAC 1183
Db 1184 TCTTGTACAGTTTGGAGGATATTAGACACTTTTGGACGATTGACGAGCTTGAAC 1243
QY 1184 CTTACAGCGCAATTAAGAGATGGATCCGTCGCGATGGAATGCCCTTCCAGATATAT 1243
Db 1244 CTTACATCTGCAATTAAGAGATGGAATTCATCAGAGATAGAACACCTTCCAGATATAT 1303
QY 1244 GAAAGAGTGTACATGATGTTTATCACACCGTAAATGAATGGCTCGAGTGGCAGAA 1303
Db 1304 GAAATGTGTACATGATGTTTATCACTATGTTTGAAGCTGTAATGAACCTGACACGAGAGCGGAGAA 1363
QY 1304 GGCTCAAGCGGAGACACGCTCAACTATGCAAGACAGCTTGGGAGGCGTGTGTTGATTC 1363
Db 1364 GACTCAAGGAGAACACTCTCACTATGTTGCAAGGCTTGGGAGGCTTATTTGATTC 1423
QY 1364 GTATATGAGGAGCAAGTGGATCGCAGCTGGTTATCTGCCACGTTTGGAGGATCTT 1423
Db 1424 ATATATGAAGAAGCAAAATGGATCTCTAATGGTTATCTGCCAATGTTTGAAGAGTACCA 1483
QY 1424 GGAGACGGGAAGTTAGCTCTGCTCATCGCCATGCCATCGCAGCTGCACCCATCTGACGTT 1483
Db 1484 TGAGATGGGAAGTGAAGCTCTGCATATCGCGTAGCAATATGCAACCCATCTCCTCCTT 1543
QY 1484 GGACATCCCTTCTCTGATCATCATCTCAAGGAGTTCAGCTTCCATCGAAGCTCAATGA 1543
Db 1544 GAATGATGGCTTCTGATATCATCTTGAAGGATGATTTTCCATCCAGGTTCAATGA 1603
QY 1544 CTTGATATGATCATCTCTGATTAAGAGTGTATACAGGTTGATACAGGAGCAGAGGC 1603
Db 1604 TTTGGCATCGCTTCTCTGATTAAGAGTGTATACAGGTTGATACAGGAGCAGAGGC 1663
QY 1604 CCGTGGAGAGAGCTGCTATATCATGTTATATGAAGACATCTCTGATTAACCGGA 1663
Db 1664 TCGTGGTGAAGAGCTTCTGATATCAATGTTATATGAAGACATCTCTGATTAACCGGA 1723
QY 1664 AGAAGATGCTCTGATATCATCACTTCAATGATCAGGAGCAGCAATCAGAGATTAATTTG 1723
Db 1724 AGAAGATGCTCTCAATCATATCAATGATCAGGAGCAGCAATCAGAGATTAATTTG 1783
QY 1724 GGAGCTTCTAAGCAGACACAGTGTTCCTCATCTTCCAGAGAGCAGCAGTGTGACAT 1783
Db 1784 GGAACTTCTAAGATCCACAGCAATATTCCTAATGCTGGCCAGAGAAACATGCTTTTGACAT 1843
QY 1784 AAGCAGAGTTTGGCATCAGGTTACAGATACGAGATGGCTACAGCTTGGCAGGTTGA 1843
Db 1844 AACAGAGCTCTCCACCATCTCATATATATCGAGATGGCTTTAGTGTGCGCAACAGGA 1903
QY 1844 AACAAAGAGTTTGGTATGAGAACCGTCAATGAAACCTGTGCTTTGTAACAACTTCAA 1903
Db 1904 AACAAAAAATTTGTTATGAGAAACACCTCTGATCTATGCTTTTAACTATTAACATA 1963
QY 1904 ATCTACAAATTAATCTAGGAGTGCCTATGCGGTGTATATAGGCA 1950
Db 1964 TCCATAATAAAGCTCATAATGCTAAATTTATGGCCTTATGACATA 2010
```

## RESULT 13

```
US-09-893-820-29
; Sequence 29, Application US/09893820
; Publication No. US20040053386A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20040053386A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
```

```
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 2196
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (69)...(1949)
; OTHER INFORMATION: myrcene synthase
US-09-893-820-29
```

```
Query Match 53.2%; Score 1071.6; DB 13; Length 2196;
Best Local Similarity 75.0%; Pred. No. 1.4e-306;
Matches 1430; Conservative 0; Mismatches 441; Indels 36; Gaps 6;

QY 68 CAGTCTGCTCAGTTCTTCTCATGAGATTAAGGCTCTCGTAGAACAATCCCAACTCT 127
Db 116 CAAGTCGTTGATCAGTTCAATTCATGAACATTAAGCCTCCCTATAGAACATCCCAATCT 175
QY 128 TGGATCTGAGCGGGGAAATCCGTCGCGCATTCATAAACATGTTTGAACAGGT 187
Db 176 TGGAAATGCTAGGCGAGGAAATCTGTACGCTTCCATGAGCATCAGTTTGGCCACCGC 235
QY 188 CGCATCTACTGATCTGTACAGAGACGGTGGCAACTATCATTCACACCTGTGGAGCA 247
Db 236 TGCACCTGATGATGGTGTACAAAGCGCATAGGTGATACCATTCCTCAATATCTGGAGCA 295
QY 248 TGATTTTCATACAGTCTCTGATCTCAACGCTTATGAGGACCTGATTCACGGGAACTGTC 307
Db 296 TGATTTTCATACAGTCTC---TATCAACGCTTATGGGAACTTACAGGAACTGTC 352
QY 308 TGACAGACTTATGGGAAAGTAAAGGATATATGTTCAATTCAGTCCGCTGGAAGATGG 367
Db 353 TGAGAGATTAATTTGGAGGTAAAGAGATA---TTCAATTCATGATACCTGGATGATGG 409
QY 368 AGG-----CAATGATCTCTTCAACGACTTTTGTGTCGATGAGCTTGA 412
Db 410 AAGATTATGAGTTCCTTTATGATCTCATGCAAGCCTTTGGATAGTCGATAGCTTGA 469
QY 413 ACGTTTGGGAATCGACAGGCAATTTCAAAAAGAGATAAAACGGCACTCCGATTTATGTTAA 472
Db 470 ACGTTTGGGATAGCTAGACATTTCAAGAACGAGATAACATCAGCTCTGGATTAATGTTT 529
QY 473 CAGTTATGGAACGAAAAAGGCATTTGATGTGGAGGAGAGTGTGTGACTGACCTCAA 532
Db 530 CCGTTACTGGAGAGAAACGGCATTTGGATGTGGAGAGACAGTATTTGTACTGATCTCAA 589
QY 533 CTCAACCGCTTGGGGCTTCGCACTCCGACTACAGGATACACTGTGTTCTCAGATGT 592
Db 590 CTCACTGCGTTGGGGTTTGAACCTTCGATACACGGGTACACTGTATCTCCAGAGGT 649
QY 593 TTTGAACGTTTTTAAAGACAAAAATGGGCAATTTTCTCCACTGCCAATATTCAGATAGA 652
Db 650 TTTAAAAGCTTTTCAAGATCAAAATGGACAGTTTGTATGTCCTCCCTCC---GGTCAGACAGA 706
QY 653 GGGAGAGATTAGAGCGCTTCTCAATTTATTCAGGGCCCTCCTCGTCCCTTTCCCGGCA 712
Db 707 GGGTGAGATCAGAAAGCGTTCTTAACTATATATCGGGCTTCCCTCATTTCCCTTCCCTGTGA 766
QY 713 GAAAGTTATGATCAAGCTGAAACATTTCTTACAAAAATATTTAAGAGAAAGCCCTGCAAAA 772
Db 767 GAAAGTTATGAAAGAGCTGAAATCTTCTTCAACAAAGATATTTGAAAGAGCTCTACAAA 826
```

QY	773	GA	TTCCGGCA	TCCAGT	A	CTTT	CA	CTAG	AGAT	ACGG	ACGTTT	CT	CTGG	AAATAT	CGTTGG	CA	832													
Db	827	GA	TTCCAGT	CTCCGCT	---	CTTT	CA	CAAG	AGATAA	AGTTT	GT	TAT	GGAA	TAT	CGCTGG	CA	883													
QY	833	CAC	CAATTTCC	CCACGCTT	CGA	CCG	CA	AGCA	AGGA	ATTA	CTAC	TG	AGCGT	CTTT	GG	ACAG	CACTA- 891													
Db	884	CACA	ATTTCC	CAAGATTT	GG	AGCA	AGAA	ATTT	AC	TAG	ACAC	ACTT	GG	AAAG	CA	CACCAG	943													
QY	892	-----	AAAA	TAAG	AAAC	CCCG	CCG	AGAA	AACTTTT	TAG	AACTTT	TG	CAAA	TTT	G	CAATTT	943													
Db	944	TG	CA	TG	CGCT	CA	ATAA	AAAT	CGT	CGG	AAAG	AGCTTTT	TAG	AACTTT	TG	CAAA	TTGAGTTCAA 1003													
QY	944	TAT	ATTT	CACT	CCCTT	TAC	AG	AGAG	AGAGTTT	AA	ACAT	GT	TCCG	ATG	TTG	CGAAG	CACT 1003													
Db	1004	TAT	ATTT	AACT	CCCTT	TAC	CA	CAAA	AGGA	ATTTAC	CAAT	ATCTTTT	GG	ATG	TTG	CGAAG	CACT 1063													
QY	1004	GGG	TTCT	CC	TGAG	ATG	ACCTTT	CT	GT	CG	ACAT	CGT	CAG	CTGG	AA	TACTAC	CTTTGGCTTC 1063													
Db	1064	GG	ATTT	GCCT	TAAT	TTG	ACA	TTT	GTCT	CG	CA	TCGT	CA	TG	TG	CACTTT	GGCCTC 1123													
QY	1064	CTG	CA	TTG	CGGT	CG	AGCCT	CA	AA	TTCT	GT	GA	TT	CAG	ACT	CGGCTTT	AC	CAAGATGCTCA 1123												
Db	1124	TTG	TA	TTG	CCCA	ATT	G	ACCC	CA	AAAC	CA	TTCT	CG	ACT	TAG	CGT	TCGCC	CAAAATGTGTCA 1183												
QY	1124	TC	TTAT	CA	CGGTT	CTT	C	AC	CA	CA	TG	TAC	AGCGT	CTT	CG	CA	CA	CACTAC	CAGCTGGAAC 1183											
Db	1184	TC	TTG	CA	CAGTTT	GG	AC	CA	TATTT	TAC	GA	CACTTTT	GG	AA	CG	ATTT	G	AC	CGAGCTTGBACT 1243											
QY	1184	CTT	CA	AG	CGCA	CA	ATT	AA	GAT	GG	GA	TCCGT	CG	CG	AT	GG	AA	TG	CTTCC	GAATATAT 1243										
Db	1244	CTT	C	ACA	TCT	CT	CA	ATT	AA	GAT	GG	AA	TT	CA	T	CAG	AT	GA	CA	CACCTTCC	AGAATATAT 1303									
QY	1244	GAA	AG	AGT	GT	CA	CA	TG	TTT	TAT	CA	CG	TA	AA	T	G	AA	T	GC	AGT	TGG	CGAGAA 1303								
Db	1304	GAA	AT	GT	GT	CA	CGT	CGT	GT	TT	G	AA	CT	G	T	AA	AT	G	AA	CT	G	AC	AG	AGCGCGAGAA 1363						
QY	1304	GGC	TT	CA	AGG	CGG	AG	CA	CGCT	CA	ACT	AT	G	CA	AG	CA	AGCGCTT	TGG	AG	CGCGT	GT	TTT	TG	ATTC 1363						
Db	1364	GACT	CA	AGG	GAG	AA	CA	CACT	CT	CA	ACT	AT	G	T	TC	GA	AA	GGCTT	TGG	AG	CGCTT	ATTT	TG	ATTC 1423						
QY	1364	GT	AT	TG	AGG	AG	CA	AA	GT	GG	AT	CG	CA	CT	GT	TT	AT	CT	G	CC	CA	CGTTT	TG	AG	ATCTT 1423					
Db	1424	AT	AT	TG	CA	AG	AG	CA	AA	AT	GG	AT	CT	CT	AA	T	GG	TT	AT	CT	G	CA	AT	TTT	GA	AG	AT	TACCA 1483		
QY	1424	GG	AA	ACGG	AA	AGTT	AG	CT	CT	G	CT	AT	CG	CC	AT	TGG	CG	CA	CTG	CA	AA	CCCA	TTCT	G	AC	GT 1483				
Db	1484	TG	AG	AT	TGG	AA	AG	TT	AG	CT	CT	G	AT	AT	CG	GT	AG	CA	CA	AT	TG	CA	CCCA	TTCT	C	CACTTT 1543				
QY	1484	GG	CA	T	CCCTTT	CC	T	GA	T	CA	CT	CA	AG	AA	GT	TGA	CTT	CC	AT	TG	CA	AG	CT	CA	AT	G	A 1543			
Db	1544	GA	T	G	CA	T	GGCTT	CC	T	GA	TT	AC	CTTT	GA	AG	GG	AA	TT	GA	TTT	CC	AT	CC	AG	GT	TT	CA	AT	G	A 1603
QY	1544	CTT	G	AT	T	AT	CA	T	CT	TC	GA	TT	AA	G																

Db	1904	AAACAAAAATTTGGTTATGGAACACCTCTTGAATCTATGCTTTTAACTATAACCATTA	1963
Qy	1904	ATCTACAATATTAACTGAGGATGCCCTATGGGTGTATATAGGGCACA	1950
Db	1964	TCCATAATAATAGCTCAATATGCTAAATATTGGCCTTATGACATA	2010
RESULT 14			
US-10-041-007-25			
; Sequence 25, Application US/10041007			
; Publication No. US20020164736A1			
; GENERAL INFORMATION:			
; APPLICANT: Matsuda, Seiichi P.T.			
; APPLICANT: Schepmann, Hala G			
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthese			
; FILE REFERENCE: P02081US1			
; CURRENT APPLICATION NUMBER: US/10/041,007			
; CURRENT FILING DATE: 2002-01-07			
; PRIOR APPLICATION NUMBER: US 60/259,881			
; PRIOR FILING DATE: 2001-01-05			
; NUMBER OF SEQ ID NOS: 41			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 25			
; LENGTH: 2196			
; TYPE: DNA			
; ORGANISM: Abies grandis			
US-10-041-007-25			
Query Match 53.2%; Score 1071.6; DB 14; Length 2196;			
Best Local Similarity 75.0%; Pred. No. 1.4e-106;			
Matches 1430; Conservative 0; Mismatches 441; Indels 36; Gaps 6;			
Qy	68	CAGTCGTGCTCAGTTCTTCTCATGAGATTAAGGCTCTCCGTAGAACAAATCCAACTCT	127
Db	116	CAAGTCGTGTGATCAGTTTCAATTCATGAACATAAGCCTCCCTATAGAACAAATCCAAATCT	175
Qy	128	TGGAATCTGAGCGCGGGGAAATCCGTCGGCATTCCATAAACATGTGTTGACAAGGCT	187
Db	176	TGGAATCGGTAGCGGAGGAAATCTGTCAGCCCTTCATGAGCAATCAGTTTGGCCACCGC	235
Qy	188	CGCATCTACTGATTTCTGTACAGACGCGTGGGCAACTATCATTCCAACCTGTGGGACGA	247
Db	236	TGCACCTGATGATGGTGTACAAAGCCATAGGTGATCACTCATTCGAATATCTGGGACGA	295
Qy	248	TGATTTTCATACAGTCTCTGATCTCAACGCCCTTATGGAGCACCTGATTAACCGGAACTGC	307
Db	296	TGATTTTCATACAGTCTC---TATCAACGCCCTTATGGGAAACCCCTCTTACCAGGAACGTGC	352
Qy	308	TGACAGACTTATTGGGGAAGTAAAGGATAATAGTTTCAATTTCAAGTCGCTGGNAGATGG	367
Db	353	TGAGAGATTAAATTGTGGAGGTAAAGAAGATA---TTCAATTTCAATGTACCTGGATGATGG	409
Qy	368	AGG-----CAATGATCTCTTCAACGACTTTTGTCTGTCGATGACGTTGA	412
Db	410	AAGATTAAATGAGTTCCTTTAATGATCTATGCAACGCCCTTTGGATAGTCGATAGCGTTGA	469
Qy	413	ACGTTTGGGAATCGACAGGCATTTCAAAAAGAGATAAAAACGGCATCTCGAATATGTTAA	472
Db	470	ACGTTTGGGATAGCTAGACATTTCAAGAACGAGATAACATCAGCTCTGGATTATGTTTT	529
Qy	473	CAGTTATTGNAACAAAAGAGCATGTGGAGGAGAGATGTTCTGTGACTGACCTCAA	532
Db	530	CCGTTACTGGGAGAAAACGCGATTTGGATTGGGAGAGACAGTATTTGTTACTGATCTCAA	589
Qy	533	CTCAACCGCCTTGGGGCTTGGAACTCTCCGACTACACGGATACACTGTGCTTCCAGATGT	592
Db	590	CTCAACTGCGTTGGGGTTTGGAACTCTTCGATTACACGGGTACACTGTATCTCCAGAGGT	649
Qy	593	TTTGAACGTTTTTAAAGACAAAAATGGGCAATTTTCTCCACTGCCCAATATTTCAGATAGA	652
Db	650	TTTAAAGCTTTTCAAGATCAAAATGACACAGTTTGTATGCTCCCC---GGTCAGACAGA	706





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:38:02 ; Search time 59 Seconds  
(without alignments)  
2959.565 Million cell updates/sec

Title: US-10-025-145a-65

Perfect score: 3251

Sequence: 1 MALLSITPLVRSCLSSSHE.....FANVETKSLVMRTVIEPVPL 618

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq 29Jan04:\*

1: \_geneseqp1980s:\*

2: \_geneseqp1990s:\*

3: \_geneseqp2000s:\*

4: \_geneseqp2001s:\*

5: \_geneseqp2002s:\*

6: \_geneseqp2003as:\*

7: \_geneseqp2003bs:\*

8: \_geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3251	100.0	618	4	AAB69390	Grand fir
2	2453	75.5	628	2	AAW85701	Pinene sy
3	2453	75.5	628	3	AAy90837	Grand fir
4	2453	75.5	628	4	AAB69371	Grand fir
5	2264	69.6	630	4	AAB69391	Grand fir
6	2200.5	67.7	627	2	AAW85700	Myrcene s
7	2200.5	67.7	627	3	AAy90842	Grand fir
8	2200.5	67.7	627	4	AAB69370	Grand fir
9	2197.5	67.6	627	2	AAW85710	Grand fir
10	2197.5	67.6	627	4	AAB69380	Grand fir
11	2197.5	67.6	627	5	ABB79395	Taxadiene
12	2137	65.7	630	4	AAB69393	Grand fir
13	2084.5	64.1	637	4	AAB69392	Grand fir
14	2031.5	62.5	637	2	AAW85702	Limone
15	2031.5	62.5	637	3	AAy90859	Grand fir
16	2031.5	62.5	637	4	AAB69372	Grand fir
17	1617.5	49.8	462	3	AAAB18115	Pinus rad
18	1611.5	49.6	462	3	AAAB18044	Pinus rad
19	1295.5	39.8	580	2	AAy06571	Delta-sel
20	1295.5	39.8	581	2	AAy06570	Delta-sel
21	1295.5	39.8	581	2	AAy06563	Grand fir
22	1295.5	39.8	581	3	AAy90854	Grand fir
23	1291.5	39.7	581	2	AAy06569	Delta-sel
24	1267.5	39.0	344	3	AAAB18043	Pinus rad
25	1250.5	38.5	577	2	AAW85704	Grand Fir

26	1250.5	38.5	577	4	AAB69374	Aab69374 Grand fir
27	1240.5	38.2	862	2	AAW31655	Aaw31655 Pacific Y
28	1240.5	38.2	862	3	AAy90852	Aay90852 Yew taxad
29	1240.5	38.2	862	5	ABB79394	Abb79394 Taxadiene
30	1236.5	38.0	862	6	ABU09782	Abu09782 Pacific Y
31	1231	37.9	593	2	AAy06572	Aay06572 Gamma-hum
32	1231	37.9	593	2	AAy06572	Aay06572 Gamma-hum
33	1231	37.9	593	3	AAy90855	Aay90855 Grand fir
34	1230	37.8	593	3	AAy06573	Aay06573 Gamma-hum
35	1230	37.8	593	2	AAy06574	Aay06574 Gamma-hum
36	1188.5	36.6	817	2	AAy06566	Aay06566 Grand fir
37	1188.5	36.6	817	5	ABB79391	Abb79391 Taxadiene
38	1187	36.5	782	2	AAW85703	Aaw85703 Grand fir
39	1187	36.5	782	3	AAy90853	Aay90853 Grand fir
40	1187	36.5	782	4	AAB69373	Aab69373 Grand fir
41	1187	36.5	782	5	ABB79392	Abb79392 Taxadiene
42	1187	36.5	817	2	AAy06567	Aay06567 E-alpha-b
43	1187	36.5	817	2	AAy06568	Aay06568 E-alpha-b
44	1187	36.5	817	2	AAy06562	Aay06562 Grand fir
45	1187	36.5	817	6	ABU09781	Abu09781 White fir

## ALIGNMENTS

## RESULT 1

AAAB69390

ID AAB69390 standard; protein; 618 AA.

XX AC AAB69390;

XX DT 30-APR-2001 (first entry)

XX DE Grand fir monoterpene synthase protein fragment SEQ ID NO: 65.

XX KW Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;

XX KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;

XX KW terpinolene synthase; insect resistance; nutrition.

XX OS Abies grandis.

XX PN WO200107565-A2.

XX PD 01-FEB-2001.

XX PF 24-JUL-2000; 2000WO-US020264.

XX PR 26-JUL-1999; 99US-00360545.

XX PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX PI Steele CL, Bohlmann J, Croteau RB, Phillips MA;

XX DR WPI; 2001-182782/18.

XX DR N-PSDB; AAF73411.

XX PT New nucleic acid encoding monoterpene synthases, for increasing terpene synthesis in plants, e.g. for increasing resistance to pests or for treatment of cancer.

XX PS Claim 7; Page 150-151; 175pp; English.

XX CC The present invention provides the protein and coding sequences of monoterpene synthases from the grand fir. These include (-)-camphene synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase and pinene synthase. The sequences can be used to produce transgenic plants expressing high levels of the enzymes, resulting in levels which are useful in protecting against and treating cancers, and to confer insect resistance on plants

XX SQ Sequence 618 AA;

Query Match 100.0%; Score 3251; DB 4; Length 618;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-295; Indels 0; Gaps 0;  
 Matches 618; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALLSITPLVSRCLSSSHEIKALRRTIPTLGICRPGKSVAHINNCLTSVASTDSVQR 60  
 DB 1 MALLSITPLVSRCLSSSHEIKALRRTIPTLGICRPGKSVAHINNCLTSVASTDSVQR 60

QY 61 VGNHNSLWDDDDFIQSLSTPYGADPYRERADRLIGEVKDIMFNFKSLDGGNDLLQRL 120  
 DB 61 VGNHNSLWDDDDFIQSLSTPYGADPYRERADRLIGEVKDIMFNFKSLDGGNDLLQRL 120

QY 121 LVDDVERLGDHDFKKEIKTALDYVNSYWNKEGIGCGRESVWTDLNSALGLRTLRHGY 180  
 DB 121 LVDDVERLGDHDFKKEIKTALDYVNSYWNKEGIGCGRESVWTDLNSALGLRTLRHGY 180

QY 181 TVSSDVLNFKDNGQFSSSTANIQIEGIRGVNLFRASIVAPFGKVMDEATFTSKYL 240  
 DB 181 TVSSDVLNFKDNGQFSSSTANIQIEGIRGVNLFRASIVAPFGKVMDEATFTSKYL 240

QY 241 REALQKIPASSILSLRDLVLEYGWHTNLPRLRNMYDMVFGOHTKNKNAEKLLELAKL 300  
 DB 241 REALQKIPASSILSLRDLVLEYGWHTNLPRLRNMYDMVFGOHTKNKNAEKLLELAKL 300

QY 301 EFNIFHSLQRELKHVSRWVKDGSPEMTFCRRHVEYVALASCIAPFQHSGRFLGFTK 360  
 DB 301 EFNIFHSLQRELKHVSRWVKDGSPEMTFCRRHVEYVALASCIAPFQHSGRFLGFTK 360

QY 361 MSHLITVLDDMDYVFGVDELELFTATIKWDPNAMECLPEYMKGVVMYVHTVNMARV 420  
 DB 361 MSHLITVLDDMDYVFGVDELELFTATIKWDPNAMECLPEYMKGVVMYVHTVNMARV 420

QY 421 AEQAQRDTLNYARQAEACFDSYMOEAKWIATGYLPTFEYLENGKVSAAHRCALQPI 480  
 DB 421 AEQAQRDTLNYARQAEACFDSYMOEAKWIATGYLPTFEYLENGKVSAAHRCALQPI 480

QY 481 LTLDIPFPDHLKEVDPPSKNDLICIILRLRGDTRCYKADRAEGEASSISCYMKDNGP 540  
 DB 481 LTLDIPFPDHLKEVDPPSKNDLICIILRLRGDTRCYKADRAEGEASSISCYMKDNGP 540

QY 541 LTEDALNHNFMIRDAIRELNWELLKPDNSVPITSKKHAFDISRVVHHGYRVDGYSFA 600  
 DB 541 LTEDALNHNFMIRDAIRELNWELLKPDNSVPITSKKHAFDISRVVHHGYRVDGYSFA 600

QY 601 NVETKSLVMRTVIEPVL 618  
 DB 601 NVETKSLVMRTVIEPVL 618

RESULT 2  
 AAW85701 standard; protein; 628 AA.  
 AC AAW85701;  
 AC AC  
 DT 27-SEP-1999 (first entry)  
 DE Pinene synthase of grand fir.  
 DE Pinene synthase of grand fir.  
 KW Myrcene synthase; limonene synthase; pinene synthase; flavour; aroma;  
 KW defense; plant seed; oil; meal.  
 XX Abies grandis.  
 XX OS  
 XX PN W09902030-A1.  
 XX PD 21-JAN-1999.  
 XX PF 10-JUL-1998; 98WO-US014528.  
 XX PR 11-JUL-1997; 97US-0052249P.  
 XX PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX Bohlmann J, Steele CL, Croteau RB;  
 PI WPI; 1999-120396/10.  
 XX N-PSDB; AAX08644.  
 PT New isolated gymnosperm monoterpene synthase DNA - obtained from Grand  
 PT fir (Abies grandis), used to provide plants with modified production of  
 PT monoterpenes, e.g. myrcene, limonene or pinene.  
 XX Claim 13; Page 77-79; 121pp; English.  
 XX Nucleotide sequences encoding myrcene synthase, limonene synthase and  
 CC pinene synthase from Grand fir may be incorporated into any organism  
 CC (e.g. intact plant, animal, microbe), or derived cell culture that  
 CC produces geranyl diphosphate for the production of the aforementioned  
 CC enzymes or their products. The sequences when expressed in transfected  
 CC cells may also be used for the production or modification of flavour and  
 CC aroma properties, improvement of defense capability, and the alteration  
 CC of other ecological interactions mediated by myrcene, limonene, pinene,  
 CC or their derivatives. In particular they can be used for the production  
 CC of plant seeds for the extraction of oil or meal  
 XX SQ Sequence 628 AA;

Query Match 75.5%; Score 2453; DB 2; Length 628;  
 Best Local Similarity 75.3%; Pred. No. 1.9e-220;  
 Matches 475; Conservative 59; Mismatches 81; Indels 16; Gaps 6;

QY 1 MALLSITPLVSRCLSSSHEIKALRRTIPTLGICRPGKSVAHINNCLTSVASTD 55  
 DB 1 MALVSTAPLASKSLKHLKSLISSSHEIKALRRTIPTLGICRPGKSVAHINNCLTSVASTD 60

QY 56 SVQRRVGNVHNSLWDDDDFIQSLSTPYGADPYRERADRLIGEVKDIMFNFKSLDGG-- 112  
 DB 61 GVRRMGDHSNLDWDDVFIQSL-PTAYEKSYLEAEKLGIGEVKN-MFNSMSLEDGELMS 118

QY 113 --NDLLQRLVLLVDDVERLGDHDFKKEIKTALDYVNSYWNKEGIGCGRESVWTDLNSAL 170  
 DB 119 PLNDLLQRLVLLVDDVERLGDHDFKKEIKTALDYVNSYWNKEGIGCGRESVWTDLNSAL 178

QY 171 GLTRLRLHGYTVSSDVLNFKDNGQFSSSTANIQIEGIRGVNLFRASIVAPFGKVM 230  
 DB 179 GLTRLRLHGYTVSSDVLNFKDNGQFSSSTANIQIEGIRGVNLFRASIVAPFGKVM 238

QY 231 EAETESTKYLREALQKIPASSILSLRDLVLEYGWHTNLPRLRNMYDMVFGOHTKNKNA 290  
 DB 239 EAETESTKYLREALQKIPASSILSLRDLVLEYGWHTNLPRLRNMYDMVFGOHTKNKNA 297

QY 291 ---AEKILLELAKLEFNIHSLQRELKHVSRWVKDGSPEMTFCRRHVEYVALASCIAP 347  
 DB 298 YVKSXKLELAKLEFNIHSLQRELKHVSRWVKDGSPEMTFCRRHVEYVALASCIAP 357

QY 348 EPOHSGFRGLGFTKMSHLITVLDDMDYVFGVDELELFTATIKWDPNAMECLPEYMKGV 407  
 DB 358 EPOHSGFRGLGFTKMSHLITVLDDMDYVFGVDELELFTATIKWDPNAMECLPEYMKGV 417

QY 408 MMVYHTVNMARVAEKAQGRDTLNYARQAEACFDSYMOEAKWIATGYLPTFEYLENGK 467  
 DB 418 IAVDTVNMARVAEKAQGRDTLNYARQAEACFDSYMOEAKWIATGYLPTFEYLENGK 477

QY 468 VSSAHRPCALQPIILTDIPFPDHLKEVDPPSKNDLICIILRLRGDTRCYKADRAEGE 527  
 DB 478 VSCGHRISALQPIILTDIPFPDHLKEVDPPSKNDLICIILRLRGDTRCYKADRAEGE 537

QY 528 ASSISCYMKDNGPGLTEEDALNHNFMIRDAIRELNWELLKPDNSVPITSKKHAFDISRVV 587  
 DB 538 ASSISCYMKDNGPGLTEEDALNHNFMIRDAIRELNWELLKPDNSVPITSKKHAFDISRVV 597

QY 588 HHGYRVDGYSFANVETKSLVMRTVIEPVL 618  
 DB 598 HHGYRVDGYSFANVETKSLVMRTVIEPVL 628



DR WPI: 2001-182782/18.  
 DR N-PSDB; AAF73372.  
 XX  
 PT New nucleic acid encoding monoterpene synthases, for increasing terpene  
 PT synthesis in plants, e.g. for increasing resistance to pests or for  
 PT treatment of cancer.  
 XX  
 XX  
 PS Claim 66; Page 110-112; 175pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of  
 CC monoterpene synthases from the grand fir. These include (-)-camphene  
 CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 CC limonene, (-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 CC and pinene synthase. The sequences can be used to produce transgenic  
 CC plants expressing high levels of the enzymes, resulting in levels which  
 CC are useful in protecting against and treating cancers, and to confer  
 CC insect resistance on plants  
 XX  
 SQ Sequence 628 AA;  
 Query Match 75.5%; Score 2453; DB 4; Length 628;  
 Best Local Similarity 75.3%; Pred. No. 1.9e-220;  
 Matches 475; Conservative 59; Mismatches 81; Indels 16; Gaps 6;  
 QY 1 MALLSITPLVSRSL-----SSSHEIKALRRITPTLIGICRPGKSVASHNCLTSTVASTD 55  
 DB 1 MALVSTAPLASKCLHKLSSITHEIKALSTPTALGMSRRGKSIPTSISSSTVVTTDD 60  
 QY 56 SVQRRVGNVHNSLWDDDFIQSLSTPYGADPYRERADRLIGEVKDMFNFKSLDGG-- 112  
 DB 61 GVRRMGDFHNSLWDDDFIQSL-PTAYEKSYLEAEKLGIVKN-MFNSMSLEDGLMS 118  
 QY 113 --NDLQRLLLVDDVERLIGIDRHKKEIKTALDYVNSYWNKEGIGCGRESVVDLNLSTAL 170  
 DB 119 PLNDLQRLVDDVERLIGIDRHKKEIKTALDYVNSYWNKEGIGCGRESVVDLNLSTAL 178  
 QY 171 GLTRLRHGYTVSSDVLNVFKNGQPSFNTAQIEGEIRGVNLFRASIVAPFGKVM 230  
 DB 179 GLTRLRHGYTVSSDVLNVFKNGQPSFNTAQIEGEIRGVNLFRASIVAPFGKIMD 238  
 QY 231 EATFTSKYLREALQKIPASSLSLEIRDLVLEVGWHTNLPRLRNMTDVFQGHYTKKNA 290  
 DB 239 EATFTSKYLREALQKIPASSLSLEIRDLVLEVGWHTNLPRLRNMTDVFQGHYTKKNA 297  
 QY 291 ---AEKLELAKLEFNFHSLQERELKHVSRRWKKSGSPWTCRHRHVEYVALASCIAP 347  
 DB 298 YVSKKLELAKLEFNFHSLQERELKHVSRRWKKSGSPWTCRHRHVEYVALASCIAP 357  
 QY 348 EPOHSGFRLGFTKMSHLITVLDMDYDVGTVDELELFTATIKRWDPPSAMECLPEYMGVY 407  
 DB 358 EPOHSGFRLGFTKMSHLITVLDMDYDVGTVDELELFTATIKRWDPPSAMECLPEYMGVY 417  
 QY 408 MMVHTVNEARVAEKAQGRDNLNARQAEACPDYSMOBAKNIATGYLPTFBEYENGK 467  
 DB 418 IAVYDTVNEAREAEAEAGRDNLNARQAEACPDYSMOBAKNIATGYLPTFBEYENGK 477  
 QY 468 VSSAHRPCALPILTLDPDHLTKVDPPSKNDLILCIILRGDTRCYKADARGEE 527  
 DB 478 VSCGHRISALPILTLDPDHLTKVDPPSKNDLILCIILRGDTRCYKADARGEE 537  
 QY 528 ASSISCYMKONPGITEALNHNFMIRDAIRELNWELLKPDNSVPITSKHAFDISRVW 587  
 DB 538 ASSISCYMKONPGITEALNHNFMIRDAIRELNWELLKPDNSVPITSKHAFDISRVW 597  
 QY 588 HHGYRYPDGYSFANVETKSLVMTVIEPVL 618  
 DB 598 HHGYRYPDGYSFANVETKSLVMTVIEPVL 628

RESULT 5  
 AAB69391  
 ID AAB69391 standard; protein; 630 AA.  
 XX

AC AAB69391;  
 XX  
 DT 30-APR-2001 (first entry)  
 XX  
 XX Grand fir monoterpene synthase protein fragment SEQ ID NO: 67.  
 DE  
 XX Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
 KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
 KW terpinolene synthase; insect resistance; nutrition.  
 XX  
 XX Abies grandis.  
 OS  
 XX WO200107565-A2.  
 PN  
 XX 01-FEB-2001.  
 PD  
 XX 24-JUL-2000; 2000WO-US020264.  
 XX  
 XX 26-JUL-1999; 99US-00360545.  
 XX  
 XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 PA  
 XX Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
 PI  
 XX WPI: 2001-182782/18.  
 DR  
 XX N-PSDB; AAF73412.  
 DR  
 XX  
 PT New nucleic acid encoding monoterpene synthases, for increasing terpene  
 PT synthesis in plants, e.g. for increasing resistance to pests or for  
 PT treatment of cancer.  
 PT  
 XX  
 PS Claim 12; Page 154-156; 175pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of  
 CC monoterpene synthases from the grand fir. These include (-)-camphene  
 CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 CC limonene, (-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 CC and pinene synthase. The sequences can be used to produce transgenic  
 CC plants expressing high levels of the enzymes, resulting in levels which  
 CC are useful in protecting against and treating cancers, and to confer  
 CC insect resistance on plants  
 XX  
 SQ Sequence 630 AA;  
 Query Match 69.6%; Score 2264; DB 4; Length 630;  
 Best Local Similarity 68.3%; Pred. No. 1e-202;  
 Matches 435; Conservative 78; Mismatches 98; Indels 26; Gaps 9;  
 QY 1 MALLSITPLVSRSL-----SSSHEIKALRRITPTLIGICRPGKSVASHNCLTSTVAST 54  
 DB 1 MALVSSAP---KSLHKLIRSTHHEIKPLRRITPTLIGICRPGKSVASHNCLTSTVAST 57  
 QY 55 DSQRRVGNVHNSLWDDDFIQSLSTPYGADPYRERADRLIGEVKDMFNFKSLDGG-- 112  
 DB 58 DGLQRRIGDYHNSLWDDDFIQSL-STPYGEPSPRERAEKLGIVKE-MFNSMSEDESM 115  
 QY 113 ---NDLQRLLLVDDVERLIGIDRHKKEIKTALDYVNSYWNKEGIGCGRESVVDLNLSTAL 169  
 DB 116 SPLNDLIERLWVDSVERLIGIDRHKKEIKSALDYVNSYWNKEGIGCGRDSVFPDNLSTA 175  
 QY 170 LGURLRLHGYTVSSDVLNVFKNGQPSFNTAQIEGEIRGVNLFRASIVAPFGKVM 229  
 DB 176 SGFTLRLHGYTVSSDVLNVFKNGQPSFNTAQIEGEIRGVNLFRASIVAPFGKVM 234  
 QY 230 DEATFTSKYLREALQKIPASSLSLEIRDLVLEVGWHTNLPRLRNMTDVFQGHYTKKNA 285  
 DB 235 EAEIFSSRYLKEAVQKIPVSS-LSQEDYTLVGWHTNLPRLRNMTDVFQGHYTKKNA 293  
 QY 286 KKKNA-----AEKLELAKLEFNFHSLQERELKHVSRRWKKSGSPWTCRHRHVEYVAL 341  
 DB 294 KKKRTQYLDSEKLELAKLEFNFHSLQERELKHVSRRWKKSGSPWTCRHRHVEYVAL 353  
 QY 342 ASCTAFEPQHSFRLGFTKMSHLITVLDMDYDVGTVDELELFTATIKRWDPPSAMECLPE 401

Db 354 SSCATEPKHSAFRLGFAKCHLITVLDIYTFGTMDIEFLNEFNRWSEKERLPE 413  
 QY 402 YMKGVYVMYHTVNMARVAEKAQGRDITLNYARQAEACFDSYMQEAKWIATGYLPTFEE 461  
 Db 414 YMKETMALYEALTDMAEAEKTOGRDITLNYARKAWEVYLDSTQEAQWIASGYLPTFEE 473  
 QY 462 YLENGKVSSAHRPCALQPTLTLDIPPPDHILKEVDPPSKNDLICIILRLGDRTRYKAD 521  
 Db 474 YLENARKVSSGHRRAALTPLLTLDVPLDVLKIDPSPFNDLASSFLRLGDRTRYKAD 533  
 QY 522 RARGEASISYCKNDPGLTEEDALNHFNMRDAIRELNWELLPKDNVPITSKKHAF 581  
 Db 534 RDRGEASISYCKNDPGLTEEDALNHFNMRDAIRELNWELLPKDNVPITSKKHAF 593  
 QY 582 DISRVVHGHYRVDGYSFANVETKSLVMRTVIEPVPL 618  
 Db 594 EITRAFPHQLYKYRDGFSVATQETKSLVRRVTVLEPVPL 630  
 RESULT 6  
 AAW85700  
 ID AAW85700 standard; protein; 627 AA.  
 AC AAW85700;  
 XX  
 DT 27-SEP-1999. (first entry)  
 XX  
 DE Myrcene synthase of grand fir.  
 XX  
 KW Myrcene synthase; limonene synthase; pinene synthase; flavour; aroma;  
 KW defense; plant seed; oil; meal.  
 XX  
 OS Abies grandis.  
 XX  
 FN WO9902030-A1.  
 XX  
 XX 21-JAN-1999.  
 PF 10-JUL-1998; 98WO-US014528.  
 PR 11-JUL-1997; 97US-0052249P.  
 XX  
 PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 XX  
 PI Bohlmann J, Steele CL, Croteau RB;  
 XX  
 DR WPI; 1999-120396/10.  
 DR N-PSDB; AAX08643.  
 XX  
 PT New isolated gymnosperm monoterpene synthase DNA - obtained from Grand  
 PT fir (Abies grandis), used to provide plants with modified production of  
 PT monoterpenes, e.g. myrcene, limonene or pinene.  
 XX  
 PS Claim 12; Page 72-74; 121pp; English.  
 XX  
 CC Nucleotide sequences encoding myrcene synthase, limonene synthase and  
 CC pinene synthase from Grand fir may be incorporated into any organism  
 CC (e.g. intact plant, animal, microbe), or derived cell culture that  
 CC produces geranyl diphosphate for the production of the aforementioned  
 CC enzymes or their products. The sequences when expressed in transfected  
 CC cells may also be used for the production or modification of flavour and  
 CC aroma properties, improvement of defense capability, and the alteration  
 CC of other ecological interactions mediated by myrcene, limonene, pinene,  
 CC or their derivatives. In particular they can be used for the production  
 CC of plant seeds for the extraction of oil or meal  
 XX  
 SQ Sequence 627 AA;

Query Match 67.7%; Score 2200.5; DB 2; Length 627;  
 Best Local Similarity 68.1%; Pred. No. 9.1e-197;  
 Matches 427; Conservative 77; Mismatches 106; Indels 17; Gaps 7;

QY 1 MALLSITPLVRSCL-----SSSHEIKALRRRTPTLIGICRPGKSVAHNSINMCLITSVASTD 55  
 Db 1 MALVSIPLASKSCLRKSLSIHHEHKPPYRTIPNLGMERRKSVTPSPMSISLATAAPDD 60  
 QY 56 SVORVGNVHNLWDDDFIOSLISITPYGAPDYRERADRLIGEVDIMFNFKSLIEDGG--- 112  
 Db 61 GVOIRIGDHYHNLWDDDFIOSL-STPYGEPYQOERAEERLIVEVKKI-FNSMYLDDGRLMS 118  
 QY 113 --NDLLQRLLLVDDVERLIGIDRHKKEIKTALDYNSYNWKEGICGGRSVVTDLNSTAL 170  
 Db 119 SFNDLMQRLWIVDSVERLIGIAHFNKETSALDYFRYWEENGICGGRDUSIVTDLNSTAL 178  
 QY 171 GURTLRLHGYTVSSDVLNVFKDKNGQFSSTANIQIEGEIRGVNLFRASLVAFPGKVM 230  
 Db 179 GFRTLRLHGYTVSPVLKAFQDQNGQFVCS PG-QTEGEIRSVNLRYASLSLIAFPGEK 237  
 QY 231 EAETESTKYREALOKIPASSILSLEIRDVLEYGYHNTNLPRLBARNYMDVFGQHTK---N 287  
 Db 238 EAEIFSTRYLKEALOKIPVSA--LSQEIKFVMEYGYHNTNLPRLBARNYMDVFGQHTK---N 296  
 QY 288 KNAEKLLELAKLEFNIFHSLQERELKHVSRWVKDGSPEMTFCRRHVEYVALASCIAP 347  
 Db 297 KNAEKLLELAKLEFNIFNSLQELQVLLRWKESDLPKLTFAHRRHVEFTLASCIAP 356  
 QY 348 EPQHSGLRGLFTMGHLITVLDMDYDVGTVDELELFTATIKRWDPSPAMECLPEYMKGY 407  
 Db 357 DPKHSAPRLGFAKCHLVTLDDIYDTFTGTDLELFTSAIKRWNSSEIEHLPEYMKCVY 416  
 QY 408 MMVYHTVNEARVAEKAQGRDITLNYARQAEACFDSYMQEAKWIATGYLPTFEEYLENGK 467  
 Db 417 MVFTVNELTREAEKTOGRNTLNYRKAEAYFDSYMEAEKWIANGYLPMEFEEYHENGK 476  
 QY 468 VSSAHRPCALQPIITLDIPFPDHILKEVDFFSKNDLICIILRLGDRTRYKADRARGE 527  
 Db 477 VSSAYRVATLQPIITLNAWLPDYILKIDFDFSRFNDLASSFLRLGDRTRYKADRARGE 536  
 QY 528 ASSISYCKNDPGLTEEDALNHFNMRDAIRELNWELLPKDNVPITSKKHAFDISRVW 587  
 Db 537 ASSISYCKNDPGLTEEDALNHFNMRDAIRELNWELLPKDNVPITSKKHAFDISRVW 596  
 QY 588 HHGVRDGYSGFANVETKSLVMRTVIE 614  
 Db 597 HLLYIRDGFSVANKETKLVNLTLE 623  
 RESULT 7  
 AAY90842  
 ID AAY90842 standard; protein; 627 AA.  
 AC AAY90842;  
 XX  
 DT 25-AUG-2000 (first entry)  
 XX  
 DE Grand fir myrcene synthase protein sequence SEQ ID NO:30.  
 XX  
 KW Synthese; protein co-ordinate data; active site; modification; terpenoid;  
 KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;  
 KW isoprenoid; breeding programme; fragrance; flavour; pheromone;  
 KW defensive agent; pigment; antitumour; steroid hormone;  
 KW signal transduction pathway; bile acid; affinity purification;  
 KW photoreceptor; enzymatic synthesis; nutrient supplement;  
 XX  
 OS Abies grandis.  
 XX  
 FN WO200017327-A2.  
 XX  
 PD 30-MAR-2000.  
 XX  
 PF 17-SEP-1999; 99WO-US021419.  
 XX  
 PR 18-SEP-1998; 98US-0100993P.  
 PR 22-APR-1999; 99US-0130628P.

PR 23-AUG-1999; 99US-0150262P.  
 XX (KENT ) UNIV KENTUCKY RES DEPT.  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 PA  
 PI  
 PI  
 PI  
 XX Chapel J, Manna KR, Noel JP, Starks CM;  
 XX WPI; 2000-292839/25.  
 DR N-PSDB; AAA38927.  
 DR  
 XX Novel terpene synthase enzymes, useful for producing terpene  
 PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known  
 PT enzymes by specific amino acid alterations.  
 XX  
 XX Claim 117; Page 390-392; 450pp; English.  
 XX  
 CC The present invention describes an isolated terpene synthase (I)  
 CC comprising a region with at least 20% identity to region 265-535 of a 548  
 CC amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha  
 CC carbon atoms (alphaC) that have interatomic distances, between each  
 CC other, within tabulated ranges, have a centre point (within a sphere of  
 CC radius 2.3 Angstrom ) within tabulated ranges, and have an ordered  
 CC arrangement of R groups (defining aa side chains), excluding specific  
 CC tabulated arrangements (tables given in the specification). (I), and  
 CC related enzymes, are used to produce a wide range of terpenoids (e.g.  
 CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,  
 CC flavours, pheromones, defensive agents, pigments, antitumor agents,  
 CC components of signal transduction pathways, precursors of steroid  
 CC hormones and bile acids, as photoreceptors and as co-factor side chains.  
 CC Some synthases with little or no catalytic activity (and nucleic acids  
 CC encoding them) are used as controls in the analysis of products formed by  
 CC enzymatic synthesis; as nutrient supplements; for affinity purification  
 CC of isoprenoids; or to develop immunological reagents or nucleic acids for  
 CC monitoring expression of terpene synthase or inheritance of the gene in  
 CC plant breeding programs. The new synthases may produce novel terpene  
 CC products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent  
 CC sequences used in the exemplification of the present invention  
 XX  
 XX SQ Sequence 627 AA;  
 Query Match 67.7%; Score 2200.5; DB 3; Length 627;  
 Best Local Similarity 68.1%; Pred. No. 9.1e-197;  
 Matches 427; Conservative 77; Mismatches 106; Indels 17; Gaps 7;  
 QY 1 MALLSITPLVSRCL-----SSSHEIKALRRTTTLGICRPGKSVASHNMLCTSVASTD 55  
 DB 1 MALVSIPLASKSCLRKSLISSIHEHKPPYRTIPNLGMRGRGKSVTPSMISLATAAPDD 60  
 QY 56 SVQRRVGNVHNSLWDDDFIQSLSTPYGAPDYRERADRLICEVKDIFNFKSLDGG--- 112  
 DB 61 GVQRRIGDYSNHWDDFIQSL-STPYGEPYQERAEERLIVEVKKI-FNSMYLDDGR LMS 118  
 QY 113 --NDLIQRLLLVDDVERGLIDRHFKKIKTALDYVNSYWNKEGIGCGRESVVTDLNSTAL 170  
 DB 119 SFNDLMQRLWIVDSVERLGARFKHFKBITGALDYVFRYWEENGIGCGRDSIVTDLNSTAL 178  
 QY 171 GLATRLRHGTVSDVLNVKXNGQSSNTANIQIEIRGVNLPRASIVARPGEKVMVD 230  
 DB 179 GFTRLRHGTVSPEVLKAFQDQNGQVFCSPG-QTEGEIRSVNLVPRASIAFFGEKVM 237  
 QY 231 EAEFTFSKYLREALQKIPASSIISLEIRDLVLEYGWHNLPRLPEARNTMDVFGQHTK---N 287  
 DB 238 EAEFTFSKYLKEALQKIPVSA-LSQEKVMEYGWHTNLPRLPEARNTMDVLEKDTSAWLN 296  
 QY 288 KNAEKLLELAKLFNFIHSLQERELKHVSRWKKDSPEWTFCHRHVVEYALASCIAP 347  
 DB 297 KNAGKLELAKLFNFIHSLQERELKHVSRWKKDSPEWTFCHRHVVEYALASCIAP 356  
 QY 348 EPQHSFRLGFTKMSHLITVLDDMDYDVGTVDELELFTATIKRWDPPAMECLPYMKGVY 407  
 DB 357 DKHSAPFLGPAKCHIVTLDLDYDIFGIIDELELFTSAIKRWNSEIEHLPYMKGVY 416  
 QY 408 MMVYHTVNEARVAEKAQGRDTLNYARQAEACFDYSNQAEAKIATGYLPTTFEYLENGK 467

DB 417 MVVFETVNELTREAEKTOGRNTLNVYRKAWAYFDSYMEAEKWSNGYLPMEFEYHENGK 476  
 QY 468 VSSAHRPCALQPIILTDIPFDPHILKEVDPPSKLNDLICILRLRGDTTCYKADRARGE 527  
 DB 477 VSSAYRVATLQPIILTNLAWLPDYILKGIDPPFRFNDLASSFURLRGDTTCYKADRGE 536  
 QY 528 ASSISCYMKDNPGLTTEEDALNHNFMIRDAIRLNWELLKPNOSVPTSKKHAFDISRVW 587  
 DB 537 ASCISCYMKDNPGSTEEEDALNHNAMVNDIILKELNELLERSNDNIPMLAKKHAFDITRAL 596  
 QY 588 HHGYRYRDGYSPANVETKSLVMRTVIE 614  
 DB 597 HHLYIYRDGFSVANKETKKLVMETLLE 623  
 RESULT 8  
 AAB69370  
 ID AAB69370 standard; protein; 627 AA.  
 XX  
 XX AAB69370;  
 AC  
 XX 30-APR-2001 (first entry)  
 DT  
 XX Grand fir myrcene synthase SEQ ID NO: 2.  
 DE  
 XX Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
 KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
 KW terpinolene synthase; insect resistance; nutrition.  
 XX  
 OS Abies grandis.  
 XX  
 XX WO200107565-A2.  
 PN  
 XX 01-FEB-2001.  
 PD  
 XX 24-JUL-2000; 2000WO-US020264.  
 PF  
 XX 26-JUL-1999; 99US-00360545.  
 PR  
 XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 PA Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
 PI  
 XX WPI; 2001-182782/18.  
 DR N-PSDB; AAF73371.  
 DR  
 XX New nucleic acid encoding monoterpene synthases, for increasing terpene  
 PT synthesis in plants, e.g. for increasing resistance to pests or for  
 PT treatment of cancer.  
 PT  
 XX  
 PS Claim 58; Page 106-107; 175pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of  
 CC monoterpene synthases from the grand fir. These include (-)-camphene  
 CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 CC limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 CC and pinene synthase. The sequences can be used to produce transgenic  
 CC plants expressing high levels of the enzymes, resulting in levels which  
 CC are useful in protecting against and treating cancers, and to confer  
 CC insect resistance on plants  
 XX  
 SQ Sequence 627 AA;  
 Query Match 67.7%; Score 2200.5; DB 4; Length 627;  
 Best Local Similarity 68.1%; Pred. No. 9.1e-197;  
 Matches 427; Conservative 77; Mismatches 106; Indels 17; Gaps 7;  
 QY 1 MALLSITPLVSRCL-----SSSHEIKALRRTTTLGICRPGKSVASHNMLCTSVASTD 55  
 DB 1 MALVSIPLASKSCLRKSLISSIHEHKPPYRTIPNLGMRGRGKSVTPSMISLATAAPDD 60  
 QY 56 SVQRRVGNVHNSLWDDDFIQSLSTPYGAPDYRERADRLICEVKDIFNFKSLDGG--- 112

Db 61 GYQRIIGDTHSNWDDDFIQSL-STPYGEPSSQOERAEERLIVEVKKI-FNSMYLDDGRMLS 118  
 QY 113 --NDLLQRLVDDVERLIGIDRHFKKEIKTALDYVNSYWNKEGIGCGRESVVTDLNSTAL 170  
 Db 119 SFNDLMQRLWIVDSVERLIGIARHFKNEITSALDYVRYWEENGIGCGRDSIVTDLNSTAL 178  
 QY 171 GLRTLRHGYTVSSDVLNVFKDKNGQFSSSTANIQTGEIRGVNLNFRASLVAFPGKVM 230  
 Db 179 GFRTLRHGYTVSPEVLKAFQDQNGQFVCSG-QTEGEIRSVNLNFRASLVAFPGKVM 237  
 QY 231 EAETFTSKYLREALOKIPASSILSLRDVLEGVGWHNTNLPRLAEARNYMDVFGQHTK--N 287  
 Db 238 EAEIFSTRYLKEALOKIPVSA-LSQEIFVMYEGWHNTNLPRLAEARNYMDVFGQHTK--N 296  
 QY 288 KNAEKLLELAKLFIHSLQERELKHVSRWKDQSGSPMTFCRRHRHVEYVALASCIAT 347  
 Db 297 KNAEKLLELAKLFIHSLQERELKHVSRWKDQSGSPMTFCRRHRHVEYVALASCIAT 356  
 QY 348 EPOHSGFRLGFTKMSHLITVLDMDYDVGTVDELELFTATIKRWDPSPAMECLPEYMKGVY 407  
 Db 357 DPKHSAPFLGFAKCHLVTLDDIYDTFTGIDELFTSAIKRWNSSEIEHLPEYMKGVY 416  
 QY 408 MMVYHTVMEMARVAEKAQGRDTLNVARQAEACFDSYMOEAKWIATGYLPTFEYLENGK 467  
 Db 417 MVVETVNELTREAKTQGRNTLNVKAWAYFDSYMEEAKEWISNGYLPTEYVHENGK 476  
 QY 468 VSSAHRPCALQILTDIPFPDHILKEVDFFPSKNDLICIILRLRGDTRCYKADRGEE 527  
 Db 477 VSSAYRVATLQILTNLAWLPDYILKGIDFSPRENDLASSFLRLRGDTRCYKADRGEE 536  
 QY 528 ASSISCYMKDNPGLTEEDALNHINFMIRDAIRELNWELLKPDNSVPTTSKKHAFDISRV 587  
 Db 537 ASSISCYMKDNPGLTEEDALNHINAVNDI IKELNWEILLRNDNIPMLAKKHAFDITRAL 596  
 QY 588 HHGYRDRGYSFANVETKSLVMRTVIE 614  
 Db 597 HHLIYIRDGFSVANKETKLVMTLLE 623

## RESULT 9

AAW85710  
 ID AAW85710 standard; protein; 627 AA.  
 AC AAW85710;

27-SEP-1999 (first entry)

Grand Fir monoterpene synthase clone AG3.48.

Myrcene synthase; limonene synthase; pinene synthase; flavour;  
 monoterpene synthase; aroma; defense; plant seed; oil; meal; primer; PCR;  
 probe.

Abies grandis.

WO9902030-A1.

21-JAN-1999.

10-JUL-1998; 98WO-US014528.

11-JUL-1997; 97US-0052249P.

(UNIW) UNIV WASHINGTON STATE RES FOUND.

Bohlmann J, Steele CL, Croteau RB;

WPI; 1999-120396/10.

N-PSDB; AAX08663.

New isolated gymnosperm monoterpene synthase DNA - obtained from Grand  
 fir (Abies grandis), used to provide plants with modified production of

PT monoterpenes, e.g. myrcene, limonene or pinene.  
 XX  
 PS Example 3; Page 107-109; 121pp; English.  
 CC Nucleotide sequences encoding myrcene synthase, limonene synthase and  
 CC pinene synthase from Grand fir may be incorporated into any organism  
 CC (e.g. intact plant, animal, microbe), or derived cell culture that  
 CC produces geranyl diphosphate for the production of the aforementioned  
 CC enzymes or their products. The sequences when expressed in transfected  
 CC cells may also be used for the production or modification of flavour and  
 CC aroma properties, improvement of defense capability, and the alteration  
 CC of other ecological interactions mediated by myrcene, limonene, pinene,  
 CC or their derivatives. In particular they can be used for the production  
 CC of plant seeds for the extraction of oil or meal  
 XX  
 SQ Sequence 627 AA;

Query Match 67.6%; Score 2197.5; DB 2; Length 627;  
 Best Local Similarity 68.1%; Pred. No. 1.7e-196;  
 Matches 427; Conservative 77; Mismatches 106; Indels 17; Gaps 7;

QY 1 MALLSITPLVGRSCL-----SSSHEIKALRRITPTLIGICRPGKSVAHNSINMCLTSVASTD 55  
 Db 1 MALVSIPLASKSCURKSLIISIEHKPPYRTIPNLGMRGRKSVTPSPMSISLATAAPD 60  
 QY 56 SVQRRVGNHNSLNWDDDFIQSLISTPYGADYRERADRLIGEVKDIMFNFKSLEDGG-- 112  
 Db 61 GYQRIIGDTHSNWDDDFIQSL-STHYGEPSSQOERAEERLIVEVKKI-FNSMYLDDGRMLS 118  
 QY 113 --NDLLQRLVDDVERLIGIDRHFKKEIKTALDYVNSYWNKEGIGCGRESVVTDLNSTAL 170  
 Db 119 SFNDLMQRLWIVDSVERLIGIARHFKNEITSALDYVRYWEENGIGCGRDSIVTDLNSTAL 178  
 QY 171 GLRTLRHGYTVSSDVLNVFKDKNGQFSSSTANIQTGEIRGVNLNFRASLVAFPGKVM 230  
 Db 179 GFRTLRHGYTVSPEVLKAFQDQNGQFVCSG-QTEGEIRSVNLNFRASLVAFPGKVM 237  
 QY 231 EAETFTSKYLREALOKIPASSILSLRDVLEGVGWHNTNLPRLAEARNYMDVFGQHTK--N 287  
 Db 238 EAEIFSTRYLKEALOKIPVSA-LSQEIFVMYEGWHNTNLPRLAEARNYMDVFGQHTK--N 296  
 QY 288 KNAEKLLELAKLFIHSLQERELKHVSRWKDQSGSPMTFCRRHRHVEYVALASCIAT 347  
 Db 297 KNAEKLLELAKLFIHSLQERELKHVSRWKDQSGSPMTFCRRHRHVEYVALASCIAT 356  
 QY 348 EPOHSGFRLGFTKMSHLITVLDMDYDVGTVDELELFTATIKRWDPSPAMECLPEYMKGVY 407  
 Db 357 DPKHSAPFLGFAKCHLVTLDDIYDTFTGIDELFTSAIKRWNSSEIEHLPEYMKGVY 416  
 QY 408 MMVYHTVMEMARVAEKAQGRDTLNVARQAEACFDSYMOEAKWIATGYLPTFEYLENGK 467  
 Db 417 MVVETVNELTREAKTQGRNTLNVKAWAYFDSYMEEAKEWISNGYLPTEYVHENGK 476  
 QY 468 VSSAHRPCALQILTDIPFPDHILKEVDFFPSKNDLICIILRLRGDTRCYKADRGEE 527  
 Db 477 VSSAYRVATLQILTNLAWLPDYILKGIDFSPRENDLASSFLRLRGDTRCYKADRGEE 536  
 QY 528 ASSISCYMKDNPGLTEEDALNHINFMIRDAIRELNWELLKPDNSVPTTSKKHAFDISRV 587  
 Db 537 ASSISCYMKDNPGLTEEDALNHINAVNDI IKELNWEILLRNDNIPMLAKKHAFDITRAL 596  
 QY 588 HHGYRDRGYSFANVETKSLVMRTVIE 614  
 Db 597 HHLIYIRDGFSVANKETKLVMTLLE 623

## RESULT 10

AAW85710  
 ID AAW85710 standard; protein; 627 AA.  
 AC AAW85710;

30-APR-2001 (first entry)

XX DE Grand fir monoterpene synthase protein fragment SEQ ID NO: 32.  
 XX KW Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
 XX KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
 XX KW terpinolene synthase; insect resistance; nutrition.  
 XX OS Abies grandis.  
 XX OS WO200107565-A2.  
 XX PN 01-FEB-2001.  
 XX PP 24-JUL-2000; 2000WO-US020264.  
 XX PP 26-JUL-1999; 99US-00360545.  
 XX PR (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 XX PA Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
 XX PI WPI; 2001-182782/18.  
 XX DR N-PSDB; AAF73391.  
 XX DR New nucleic acid encoding monoterpene synthases, for increasing terpene  
 XX PT synthesis in plants, e.g. for increasing resistance to pests or for  
 XX PT treatment of cancer.  
 XX PS Disclosure; Page 138-139; 175pp; English.  
 XX CC The present invention provides the protein and coding sequences of  
 XX CC monoterpene synthases from the grand fir. These include (-)-camphene  
 XX CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 XX CC limonene, (-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 XX CC and pinene synthase. The sequences can be used to produce transgenic  
 XX CC plants expressing high levels of the enzymes, resulting in levels which  
 XX CC are useful in protecting against and treating cancers, and to confer  
 XX CC insect resistance on plants  
 XX CC  
 XX CC  
 XX SQ Sequence 627 AA;

Query Match 67.6%; Score 2197.5; DB 4; Length 627;  
 Best Local Similarity 68.1%; Pred. No. 1.7e-196;  
 Matches 427; Conservative 77; Mismatches 106; Indels 17; Gaps 7;

QY 1 MALLSITPLVSRCL-----SSSHEIKALRRITPTIGICRPGKSVASHINWCLTSVASTD 55  
 DB 1 MALVSIPLASKCLRSKLSSISHEHKPPYRTIPNLGMRGRGKSVTPSMSISLATAAPDD 60  
 QY 56 SVORRVGNHNSLWDDDFIOSLISSTPYGAPDYRERADRLIGEVDIMFNFKSLEDGG-- 112  
 DB 61 GVORRGDYHNSLWDDDFIOSL-STHYGEPSYQERAEELIVEVKKI-FNSMYLDDGRMLS 118  
 QY 113 --NDLQRLLLVDDVERLGDHDFHKEIKTALDYVNSYWNKEGIGCGRESVVDLNLSTAL 170  
 DB 119 SFNDLMQRLWIVDSVERLGIARHFKNEITSALDYVFRYWEENGIGCGRDSIVTDLNLSTAL 178  
 QY 171 GLBTLRLHGYTSSDVLNVFKDKNQGPSSTANTQIEGEIRGVNLFRASLVAPEGEKVD 230  
 DB 179 GPRTLRLHGYTSSDVLNVFKDKNQGPSSTANTQIEGEIRGVNLFRASLVAPEGEKVD 237  
 QY 231 EAEFTFTKYLREALQKIPASSIISLEIRDLVLYGWHNLPRLAEARNYMDVFGQHTK---N 287  
 DB 238 EAEFTFTKYLREALQKIPASSIISLEIRDLVLYGWHNLPRLAEARNYMDVFGQHTK---N 296  
 QY 288 KNAEKLELLEAKLEFNIFHSLQERELKHVSRWKDSGPWPTFCRHHVVEYALASCIAP 347  
 DB 297 KNAGCKLELLEAKLEFNIFHSLQERELKHVSRWKDSGPWPTFCRHHVVEYALASCIAP 356  
 QY 348 EPQHSGRFLGFTKMSHLITVLDMDYDVGTVDELELFTATIKRWDPSAMECLPEYMKGVY 407  
 DB 357 DPXSAFRLGPAKCHLVTVLDDIYDFGIDELELFTSAIKRWNSSIEHLEHLPYMKCVY 416

QY 408 MMVYHTVNEWARVAEKAQGRDITLYARQAEACFDYSMQEAKWIATGYLPTPEEYLENGK 467  
 DB 417 MVVFEIVNELTREAEKTOGRNTLYVRKAEAYFDSYMEAEKXWISNGYLPPTPEEYHENGK 476  
 QY 468 VSSAHPALQIPILTDIPFPDHLKEVDFPSKLNLDLCIILRLGDTTCYKADRARGE 527  
 DB 477 VSSAYRVATLQIPILTNLAWLPDYILKXGIDFPPSFNLDASSFLRLGDTTCYKADRARGE 536  
 QY 528 ASSISCYMKDNCLTEEDALNHINFMIRDAIRELNWELLKPDNSVPTTSKKHAFDISRVW 587  
 DB 537 ASCISCYMKDNCPGSTEEDALNHINAMVNDIIRKLNWELLRSNDNIFMLAKKHAFDITRAL 596  
 QY 588 HHGYYRDGYSFANVETKSLVMETVIE 614  
 DB 597 HHLYIYRDGFSVANKETKKLVWETLLE 623

RESULT 11  
 ABB79395  
 ID ABB79395 standard; protein; 627 AA.  
 XX AC ABB79395;  
 XX DT 13-AUG-2002 (first entry)  
 XX DE Taxadiene synthase homologous protein sequence w85710.  
 XX KW Taxane; taxane synthesis; taxadiene synthase; cytostatic; anticancer;  
 XX KW enzyme.  
 XX OS Unidentified.  
 XX PN WO2002040694-A2.  
 XX PD 23-MAY-2002.  
 XX PF 16-NOV-2001; 2001WO-DK000763.  
 XX PR 17-NOV-2000; 2000DK-00001730.  
 XX PR 29-NOV-2000; 2000US-0253843P.  
 XX PA (NOVO ) NOVOZYMES AS.  
 XX PI Wind J;  
 XX PS WPI; 2002-471624/50.  
 XX CC Heterologous production of taxane analogs related compound useful as  
 XX CC anticancer agent involves cloning a full-length taxane synthesis pathway  
 XX CC from a taxan-producing organism into a taxane-resistant host cell.  
 XX CC Example 3; Fig 1; 60pp; English.  
 XX CC The present invention describes the heterologous expression of taxane  
 XX CC analogues, comprising cloning a DNA sequence comprising a taxane  
 XX CC synthesis pathway, making a DNA construct in which the DNA sequence is  
 XX CC under control of regulatory elements, introducing the DNA construct into  
 XX CC a host cell, growing the host cell to produce taxane analogues, and  
 XX CC recovering the required taxane from the culture medium. Also described:  
 XX CC (1) an isolated DNA sequence comprising the taxane synthesis pathway,  
 XX CC (preferably taxol synthesis pathway); (2) an expression vector comprising  
 XX CC the isolated DNA sequence; and (3) a host cell comprising a taxane  
 XX CC synthesis pathway derived from a taxane-producing organism, plant or  
 XX CC tree, the taxane synthesis pathway is foreign to the host cell. The  
 XX CC taxane analogues have cytostatic activity, and so can be used as  
 XX CC anticancer agents. The present sequence represents a taxadiene synthase  
 XX CC homologous protein, which is given in an example from the present  
 XX CC invention  
 XX SQ Sequence 627 AA;

Query Match 67.6%; Score 2197.5; DB 5; Length 627;  
 Best Local Similarity 68.1%; Pred. No. 1.7e-196;

Matches 427; Conservative 77; Mismatches 106; Indels 17; Gaps 7;		PI Steele CL, Bohlmann J, Croteau RB, Phillips MA;	
QY	1 MALLSITPLVSRCL-----SSSHEIKALRRITPTLIGICRPGKSVASHINMCLTSVASTD 55	XX	WPI; 2001-182782/18.
Db	1 MALVSIPLASKSLCKRSLSISSHEHKPYRTIPNLGMRRCCKSVTPSWSISLATAAPDD 60	DR	N-PSDB; AAF73421.
QY	56 SVQRVGVNHLNWDGDFRTOSLSTPYGAPDYRERADRLIGVKIMFNFKSLDGG--- 112	PT	New nucleic acid encoding monoterpene synthases, for increasing terpene synthesis in plants, e.g. for increasing resistance to pests or for treatment of cancer.
Db	61 GVQRIGDVHNSIWDGDFIQSL-STHYGEPSCYQERAEERLIVEVKI-FNSMYLDDGLRMS 118	XX	Claim 50; Page 166-167; 175pp; English.
QY	113 --NDLQRLLLVDDVERLIGIDRHFKKEIKTALDYVNSYVNEKGIGCGRESVVTDLNSTAL 170	XX	The present invention provides the protein and coding sequences of monoterpene synthases from the grand fir. These include (-)-camphene synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase and pinene synthase. The sequences can be used to produce transgenic plants expressing high levels of the enzymes, resulting in levels which are useful in protecting against and treating cancers, and to confer insect resistance on plants
Db	119 SFNDLMQRLWIVDSVERLIGIAHFKNETSALDYFRYWEENGICGGRDSIVTDLNSTAL 178	XX	Sequence 630 AA;
QY	171 GLRTRLHGTVSSDVLNVFKDKNGQFSTANIQIEGEIRGVNLNFRASLVAFPPGEKVM 230	CC	Query Match 65.7%; Score 2137; DB 4; Length 630;
Db	179 GPRTRLHGTVSPVLAFAQDQNGQFVCSPG-QTEGEIRSVNLNFRASLVAFPPGEKVM 237	CC	Best Local Similarity 67.4%; Pred. No. 8.3e-191;
QY	231 EAETESTKYLREALOKIPASSLSLSLEIRDVLEYGWHTNLPRLEARNYMDVFGQHTK---N 287	CC	Matches 427; Conservative 73; Mismatches 114; Indels 20; Gaps 9;
Db	238 EAEIFSTRYLKEALQKIPVSA-LSQEIKFVMEYGWHTNLPRLEARNYMDTLEKOTSALWN 296	CC	
QY	288 KNAAEKLELAKLEFNIHFSLOERELKHVSRRWKDGGSPMTFCRRHVEYYALASCI 347	CC	
Db	297 KNAAGKLELAKLEFNIHFSLOERELKHVSRRWKDGGSPMTFCRRHVEYYALASCI 356	CC	
QY	348 EPQHSGLRGLFTKMSHLITVLDDMYDVFVTGTVDELELFTATIKRWDPSAMECLPEYMKGY 407	CC	
Db	357 DPKHSARLGLFAKMHVLTVDLDYDTFTGTDLELFTSAIKRWNSSEIEHLPEYMKCVY 416	CC	
QY	408 MMVYHTVNEMARVAEKAQGRDTLVNARQAWACFDSYMQEAKWIATGYLPTFEYLENGK 467	CC	
Db	417 MVVFTVNELTREAKTQGRNTLVNARQAWACFDSYMQEAKWIATGYLPTFEYLENGK 476	CC	
QY	468 VSSARPCALOPILTDIPFPDHILKEVDFFPSKNDLICIILRLGDRCTCYKADRGEE 527	CC	
Db	477 VSSARVATLQPILTNLAWLPDYILKGDIFSRFNDLASSFLRLGDRCTCYKADRGEE 536	CC	
QY	528 ASSISCYMKNDPGLTEEDALNHINPMIRDAIRELNWELLKPDNSVPITTSKKHAFDISRW 587	CC	
Db	537 ASSISCYMKNDPGLTEEDALNHINPMIRDAIRELNWELLKPDNSVPITTSKKHAFDISRW 596	CC	
QY	588 HHGYRYRGGYSFANVETKSLVMRTVIE 614	CC	
Db	597 HHLIYRDGFSVANKETKKLVNETHLE 623	CC	
RESULT 12			
AAB69393			
ID	AAB69393 standard; protein; 630 AA.		
XX	AC AAB69393;		
XX	30-APR-2001 (first entry)		
DE	Grand fir monoterpene synthase SEQ ID NO: 78.		
KW	Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;		
KW	myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;		
KW	terpinolene synthase; insect resistance; nutrition.		
OS	Abies grandis.		
XX	WO200107565-A2.		
XX	01-FEB-2001.		
XX	24-JUL-2000; 2000WO-US020264.		
XX	26-JUL-1999; 99US-00360545.		
XX	(UNIW ) UNIV WASHINGTON STATE RES FOUND.		
XX	RESULT 13		
XX	AAB69392		

ID AAB69392 standard; protein; 637 AA.  
 XX AAB69392;  
 AC  
 XX  
 DT 30-APR-2001 (first entry)  
 DE Grand fir monoterpene synthase protein fragment SEQ ID NO: 69.  
 DE  
 XX Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
 KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
 KW terpinolene synthase; insect resistance; nutrition.  
 XX  
 OS Abies grandis.  
 XX  
 XX WO200107565-A2.  
 PN  
 XX  
 PD 01-FEB-2001.  
 XX  
 XX 24-JUL-2000; 2000WO-US020264.  
 PF  
 XX  
 XX 26-JUL-1999; 99US-00360545.  
 PR  
 XX  
 XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 PA  
 XX Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
 PI  
 XX WPI; 2001-182782/18.  
 DR N-PSDB; AAF73413.  
 DR  
 XX New nucleic acid encoding monoterpene synthases, for increasing terpene  
 PT synthesis in plants, e.g. for increasing resistance to pests or for  
 PT treatment of cancer.  
 PT  
 XX  
 XX Claim 54; Page 159-161; 175pp; English.  
 PS  
 XX The present invention provides the protein and coding sequences of  
 CC monoterpene synthases from the grand fir. These include (-)-camphene  
 CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 CC limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 CC and pinene synthase. The sequences can be used to produce transgenic  
 CC plants expressing high levels of the enzymes, resulting in levels which  
 CC are useful in protecting against and treating cancers, and to confer  
 CC insect resistance on plants  
 CC  
 XX Sequence 637 AA;  
 SQ  
  
 Query Match 64.1%; Score 2084.5; DB 4; Length 637;  
 Best Local Similarity 64.9%; Pred. No. 7.1e-186;  
 Matches 417; Conservative 80; Mismatches 115; Indels 31; Gaps 13;  
  
 QY 1 MALLSITPL-VSRSC-----LSSSHEIKAL--RTIPLGICRPGKSAHSINMCLTSVA 52  
 DB 1 MALLSIVLQVPSKGLKSLISSNVQKALCISTAVPLRRRQKALV--INMKLITVS 58  
 QY 53 STDS-----VQRRVGNVHSLWDDDFIQSLISTIPYAGPYRERADRLLIGEVDKIMF-- 105  
 DB 59 HRDNGGGVLQRIADHPNLWEDDFIQSL-SSEYGGSSYSERAVTVVEVKE-MFNSIP 116  
 QY 106 --KSLEDGGNDLLQRLLLVDVRLGDRHFKKEIKTALDYVNSYWNK-GIGGCRSVV 162  
 DB 117 NRELFGSQNDLLTRLWMDVSIERLGDHDFQNEIRVALDVYSYWKKEGIGGCRDSTF 176  
 QY 163 TDINSTALGRLRLHGYTVSSDVLNVFKDNGQFSSTANIQIBET-RGVNLNFRASLV 221  
 DB 177 PDINSTALALTRLHGVNVSDDVLEVPKQKHGFACPA-ILFEGQITRSVNLRYASLV 235  
 QY 222 AFPGKVMDBAETFTSKYLREALQKIPASSILSLERDVLVEYGHWTNLPRLRARNYMDVF 281  
 DB 236 AFPGKVMDBAETFTSKYLREALQKIPASSILSLERDVLVEYGHWTNLPRLRARNYMDVF 294  
 QY 282 GQHTKNKN-----AAEKLELAKLEFNIHSLQERELKHVSRWKKDGGSPMTFCFRRH 335  
 DB 295 QDSYESSNEMPYNTQKLLAKLEFNIHSLQERELKHVSRWKKDGGSPMTFCFRRH 354

QY 336 VEYVALASCIAPFOHSGPRIGFTKMSHLITVLDMDYDVGFTVDELEFATIKRWDPSA 395  
 DB 355 VEYITWASCIAPFOHSGPRIGFTKMSHLITVLDMDYDVGFTVDELEFATIKRWDLS 414  
 QY 396 MECLPEYMGVVMVYHTVNMARVAEKAQGRDTLNYARQAEACDSYMOAKWTATGY 455  
 DB 415 TKCLPEYMKAVYMDLYQCLNELAQAEKQGRDTLNYIRNAYESHFDSFMHEAKWTSSGY 474  
 QY 456 LPTFEYLENGKVSNAHRPCALQPIITLDIPPDHILKSVDPKSLNDLICILRLRGDT 515  
 DB 475 LPTFEYLENGKVSNAHRPCALQPIITLDIPPDHILKSVDPKSLNDLICILRLRGDT 534  
 QY 516 RCYKADRARGEASISCYMKONPGLTEDALNHFIMIRDAIRELNWELLKPDNSVPIT 575  
 DB 535 RCYKADRARGEASISCYMKONPGLTEDALNHFIMIRDAIRELNWELLKPDNSVPIT 594  
 QY 576 SKKHAFTDIRVWHGYRDGYSFANVETKSLVMRTVIBFVPL 618  
 DB 595 SKKHAFTDIRVWHGYRDGYSFANVETKSLVMRTVIBFVPL 637  
  
 RESULT 14  
 AAW85702  
 ID AAW85702 standard; protein; 637 AA.  
 AC  
 XX AAW85702;  
 XX  
 DT 27-SEP-1999 (first entry)  
 DE Limonene synthase of grand fir.  
 DE  
 XX Myrcene synthase; limonene synthase; pinene synthase; flavour; aroma;  
 KW defense; plant seed; oil; meal.  
 KW  
 XX Abies grandis.  
 OS  
 XX WO9902030-A1.  
 PN  
 XX 21-JAN-1999.  
 PD  
 XX 10-JUL-1998; 98WO-US014528.  
 PF  
 XX 11-JUL-1997; 97US-0052249P.  
 PR  
 XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 PA  
 XX Bohlmann J, Steele CL, Croteau RB;  
 PI  
 XX WPI; 1999-120396/10.  
 DR N-PSDB; AAX08645.  
 DR  
 XX New isolated gymnosperm monoterpene synthase DNA - obtained from Grand  
 PT fir (Abies grandis), used to provide plants with modified production of  
 PT monoterpenes, e.g. myrcene, limonene or pinene.  
 XX  
 XX Claim 14; Page 82-84; 121pp; English.  
 XX  
 CC Nucleotide sequences encoding myrcene synthase, limonene synthase and  
 CC pinene synthase from Grand fir may be incorporated into any organism  
 CC (e.g. intact plant, animal, microbe), or derived cell culture that  
 CC produces geranyl diphosphate for the production of the aforementioned  
 CC enzymes or their products. The sequences when expressed in transfect  
 CC cells may also be used for the production or modification of flavour and  
 CC aroma properties, improvement of defense capability, and the alteration  
 CC of other ecological interactions mediated by myrcene, limonene, pinene,  
 CC or their derivatives. In particular they can be used for the production  
 CC of plant seeds for the extraction of oil or meal  
 XX  
 SQ Sequence 637 AA;  
 Query Match 62.5%; Score 2031.5; DB 2; Length 637;  
 Best Local Similarity 63.6%; Pred. No. 6.6e-181;



QY 396 MECLPEYMKGVYMYHTVNEARVAEKAGQDITLNYARQAEACFDSYMOBAKMIATGY 455  
DB 415 TRWLPEYMKGVYMDLYQCINEMVEAEKTOGRDMLNYIQNAWEALPDTFMQEAkWISSY 474  
QY 456 LPTFEYLENGKVSAHRPCALQPIITLDIPPPDHILKEVDPPSKLNDLICIIILRLRGDT 515  
DB 475 LPTFEYLNKAKVSSGSRIATLQPIITLDVPLPDYILQBIIDYPSRFNELASSILRLRGDT 534  
QY 516 RCYKADRARGEAEASISCYMKDNPGLTEDALNHNFMIRDAIRELNWELLKPDNSVPI 575  
DB 535 RCYKADRARGEAEASISCYMKDHPGSIEDALNHNAMISDAIRELNWELLRPDPSKSPIS 594  
QY 576 SKKHAFDISRVWHGVRDGYSEFANVETKSLVMRTVIRPVPL 618  
DB 595 SKKHAFDITRAHHVYKYRDGYTVSNNETKNLVMTVLEPLAL 637

Search completed: July 7, 2004, 17:42:14  
Job time : 61 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:40:02 ; Search time 21 Seconds  
(without alignments)  
2830.779 Million cell updates/sec

Title: US-10-025-145A-65  
Perfect score: 3251  
Sequence: 1 MALLSITPLVSRSCSSSHE.....FANVETKSLVMRTVIEPVL 618

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	724.5	22.3	599	2 A48863	limonene cyclase -
2	717.5	22.1	554	2 S68366	(+)-delta-cadinene
3	717.5	22.1	554	2 S68365	(+)-delta-cadinene
4	689.5	21.2	591	2 H84633	probable limonene
5	669.5	20.6	520	2 A56118	vetispiradiene syn
6	656	20.2	600	2 F71434	probable limonene
7	641.5	19.7	550	2 T03714	5-epi-aristolochene
8	632	19.4	559	2 T08174	sesquiterpene cycl
9	615.5	18.9	548	2 T06266	germacrene C synth
10	614.5	18.9	548	2 T06265	germacrene C synth
11	586.5	18.0	1024	2 G71434	probable limonene
12	569	17.5	632	2 E36723	hypothetical prote
13	536	16.5	598	2 H86460	hypothetical prote
14	504.5	15.5	785	2 G96825	hypothetical prote
15	504.5	15.5	785	2 T52059	ent-kaurene syntha
16	497.5	15.3	608	2 G86443	probable terpene s
17	486.5	15.0	789	2 T09672	ent-kaurene syntha
18	483	14.9	582	2 C71424	hypothetical prote
19	481	14.8	530	2 G96588	hypothetical prote
20	437.5	13.5	598	2 T00509	probable vetispira
21	427.5	13.1	421	2 C96642	hypothetical prote
22	426	13.1	598	2 F96684	probable terpene s
23	425	13.1	350	2 C56118	vetispiradiene syn
24	423.5	13.0	801	2 T06783	ent-kaurene syntha
25	415.5	12.8	573	2 T05328	hypothetical prote
26	408	12.5	802	2 D85035	ent-kaurene synthe
27	406.5	12.5	612	2 T05331	hypothetical prote
28	405	12.5	870	2 A96637	hypothetical prote
29	404.5	12.4	823	2 T02959	kaurene synthase A

## ALIGNMENTS

### RESULT 1

A48863

limonene cyclase - spearmint

C:Species: Mentha spicata (spearmint)

C>Date: 12-May-1995 #sequence\_revision 19-May-1995 #text\_change 05-Nov-1999

C:Accession: A48863

R:Colby, S.M.; Alonso, W.R.; Katahira, E.J.; McGarvey, D.J.; Croteau, R.

J. Biol. Chem. 268, 23016-23024, 1993

A:Title: 4S-limonene synthase from the oil glands of spearmint (Mentha spicata). cDNA is

A:Reference number: A48863; MUID:94043077; PMID:8226816

A:Accession: A48863

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-599 <COL>

A:Cross-references: GB:L13459; NID:g410229; PIDN:AAC37366.1; PID:g410230

C:Superfamily: vetispiradiene synthase 1

Query Match 22.3%; Score 724.5; DB 2; Length 599;  
Best Local Similarity 31.7%; Pred. No. 4.1e-43;  
Matches 197; Conservative 114; Mismatches 249; Indels 61; Gaps 16;

Qy	13	SCLSSEHIEKALRTIPTLGICRPGKSVVAHSINNCLTSTVASTDSVQRVGNVHNSLWDD	72
Db	20	TCLQPSH-----FKSSPKLLSSTNSSRSRLRVYCSSLT-----ERRSGNYPNRWDVN	71
Qy	73	FIQSLISTFYGADPYRE-----RADRLIGEYKDIMFNPKSLDGGNDLLQRLLLVDDYER	127
Db	72	FIQSLLS-----DYKEDKHVIRASELVTLVK-----MELEK-ETDQIQQLLELIDDLQR	118
Qy	128	LGIDRHFKEIKETALD-----YVNSYMWKEGIGCGRESWVTDLSNLTALGLRTLRLHG	179
Db	119	MGLSDHFQNEFKETLSSTYLDHYYKNFPFKEE-----RDLYSTSLAFRLLRHG	168
Qy	180	YTVSSDLNVFKDKNGQFSSTANTQIEGTIRGVNLFRASLVAFPGKVMDEATFTSKY	239
Db	169	FQVAQEVDFSKNEGEFKES-----LSDTRGLQLYEASFLLTGETTLESAREFATKF	224
Qy	240	LREALQKIPASSILSLRIQVLEYGWHTNLPLEARNYMDVFGQHTKNKNAEKLLELAK	299
Db	225	LEEKVGGVGDGLLTIRAYSLDIPLHWIRKRPAPVWIE---WYRKRPDMNPVLELAI	281
Qy	300	LEFNIFHSLQERELKHSRWKDSGSP-MTFCRHRHVEYYALASCIAPFPQSGFRLGF	358
Db	282	LDLNLVQAOQEELKESFRWRNTGFVEKLPFARDRLVECYFWNTGLIEPRQHASAKIMM	341
Qy	359	TKSKHLITVDDMYDVEGTVDLELFTATIKRWDPDSAMECLPEYMKGVYMMVHTVNEA	418
Db	342	GKNALITVDDIYDVGTLLEELEQFTDLIRRWIDINSIDQLPDYMQCLFLALNNFVDDTS	401
Qy	419	RVAEKAQGRDTLNVARQAEACFDSYMOEAKVIATGYLPTFEFYLENGKSSAHRPCLAQ	478
Db	402	YDWKKEGVNVPYLRQSWVDLADKYMVEARFYGGHKPSLEEYLEN-SWQSIGSPCMLT	460

```
QY 479 PTLTLDIPFDHILKE-VDPFSLKNDLI---CIIILRGDTRCYKADRGEEASSISCY 534
Db 461 HIF---FRYVDSFTKTVSLYKHDLVRSWGSFVLELADDLGTSVEVRGDPVKSLQCY 517
QY 535 MKDNPGLTEDALNHNFMIRDAIRLNWELLKPDNSVPITSKKHAFDLSRVVWHGYYR 594
Db 518 MDGY-NASAEARKHVWLLAEVWKONAEVSKDSFGKDFGICAVDLGRMAQLMYHG 576
QY 595 DGYSFANVETKSLVMTVIEP 615
Db 577 DGHGTQHPHIIHQWMTLPEP 597

RESULT 2
S68366
(+)-delta-cadinene synthase isozyme XC14 - Gossypium arboreum
C:Species: Gossypium arboreum
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 29-Sep-1999
C:Accession: S68366
R:Chen, X.Y.; Chen, Y.; Heinstein, P.; Davison, V.J.
Arch. Biochem. Biophys. 324, 255-266, 1995
A:Title: Cloning, expression, and characterization of (+)-delta-cadinene synthase: a cat
A:Reference number: S68365; MUID:96132653; PMID:8554317
A:Accession: S68366
A:Molecule type: mRNA
A:Residues: 1-554 <CHE>
A:Cross-references: EMBL:U23205; NID:g1045313; PIDN:AAA93065.1; PID:g1045314
A:Experimental source: cultivar Nanking
C:Superfamily: vetispiradiene synthase 1
C:Keywords: phytoalexin biosynthesis

Query Match 22.1%; Score 717.5; DB 2; Length 554;
Best Local Similarity 31.9%; Pred. No. 1.1e-42;
Matches 184; Conservative 108; Mismatches 241; Indels 43; Gaps 14;

QY 51 VASTDSVQRVGVNHNWDDDFIQSLISTPYGAPD-----YRERADRLIGVKKDIFEN 104
Db 14 LSNKDEMRPKADFPQSIWGDJFL-----NCPKNIDAETEKRHQQLKEVR----- 60
QY 105 FKSLEDGNDLLQRLLLVDDVERLGDHFKKBIKTALDYVNSYWNKEGIGCGRESVVD 164
Db 61 -KMIVAPMANSTOKLAFIDSVQSLGVSYHFTKEIEDELE--NIYHN-----NDAEND 110
QY 165 INSTALGLRTLRLHGTVSSDVLNVKDKNGQSSSTANIQIEGIRGVNLFRASLVAPP 224
Db 111 LYTTSLRFLRLRHEGVNVCDFNFKDQGNFKSS-----VTSVGRGLLELYQASLYRVH 166
QY 225 GEKVMDEAETFTSKYLREALQKIPASSILSLERDVLVEYGVWHTNLPRLERNTMDVFGQH 284
Db 167 GEDILDEAISFTTNHLSLAVASLDP--LSEEVSHALKQSIIRGLPRVEARHVLVSV-QD 223
QY 285 TKKNAAEKLELAKLEFNIFHSIQRELBKHVGRWKD--SGSPMTFCRRHVEYYALAS 343
Db 224 IESHN--KVLEFAKIDFNWQLLHRELSEISRWKDKLDFQKLPYARDVVEGYFWIS 281
QY 344 CIAFEPQHSGLRGTQKMSHLITVLDMDYVFGTVDELELFTATIKWDPSAMECLPEYM 403
Db 282 GVTFEPQYSLGRKMLTKVIAMASIVDDYTSYATYELIPTYKAIERWDIKCIDELPEYM 341
QY 404 KGVMMVYHTVNEMARVAEKAQGRDTLNYARQAWACFDSYMOEAKWIAATGYLPTPEYL 463
Db 342 KPSYKALLDVVEEMVQLVAEHGQYRVEYAKNAMIRLAQSYLVEARWTLQNYKPSFEFK 401
QY 464 ENKVSAHRPCALPILTL-DIPFPDHLKEVDVFPKNDLICIIILRLRGDTRCYKADR 522
Db 342 KPSYKALLDVVEEMVQLVAEHGQYRVEYAKNAMIRLAQSYLVEARWTLQNYKPSFEFK 401
QY 464 ENKVSAHRPCALPILTL-DIPFPDHLKEVDVFPKNDLICIIILRLRGDTRCYKADR 522
Db 402 ANALPTCGYAMLATISFVGMGDIVTPTFKWAANDP-KITQASTIICRFMDVVAEHKPKH 460
QY 523 ARGEASSISCYMKDNPGLTEEDALNHNFMIRDAIRLNWELLKPDNSVPITSKKHAFD 582
Db 461 RREDDCSAIECYMEY-GVTAQAYDVFNKHVESAWKDVNKEFLKP-TEMPTEVLNLSN 518
QY 583 ISRVWHHGYYRDGYSFANVETKSLVMTVIEPVL 618
QY 519 LARVMDVLYREGDGYTVVGKAAKGGITSLIIIEPTAL 554
```

Db 519 LARVMDVLYREGDGYTVVGKAAKGGITSLIIIEPTAL 554

## RESULT 3

S68365

(+)-delta-cadinene synthase isozyme XC1 - Gossypium arboreum

C:Species: Gossypium arboreum

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 29-Sep-1999

C:Accession: S68365

R:Chen, X.Y.; Chen, Y.; Heinstein, P.; Davison, V.J.

Arch. Biochem. Biophys. 324, 255-266, 1995

A:Title: Cloning, expression, and characterization of (+)-delta-cadinene synthase: a cat

A:Reference number: S68365; MUID:96132653; PMID:8554317

A:Accession: S68365

A:Molecule type: mRNA

A:Residues: 1-554 <CHE>

A:Cross-references: EMBL:U23206; NID:g1045311; PIDN:AAA93064.1; PID:g1045312

A:Experimental source: cultivar Nanking

C:Superfamily: vetispiradiene synthase 1

C:Keywords: phytoalexin biosynthesis

Query Match 22.1%; Score 717.5; DB 2; Length 554;

Best Local Similarity 31.8%; Pred. No. 1.1e-42;

Matches 183; Conservative 107; Mismatches 243; Indels 43; Gaps 14;

QY 51 VASTDSVQRVGVNHNWDDDFIQSLISTPYGAPD-----YRERADRLIGVKKDIFEN 104

Db 14 LSNKDEMRPKADFPQSIWGDJFL-----NCPKNIDAETEKRHQQLKEVR----- 60

QY 105 FKSLEDGNDLLQRLLLVDDVERLGDHFKKBIKTALDYVNSYWNKEGIGCGRESVVD 164

Db 61 -KMIVAPMANSTOKLAFIDSVQSLGVSYHFTKEIEDELE--NIYHN-----NDAEND 110

QY 165 INSTALGLRTLRLHGTVSSDVLNVKDKNGQSSSTANIQIEGIRGVNLFRASLVAPP 224

Db 111 LYTTSLRFLRLRHEGVNVCDFNFKDQGNFKSS-----VTSVGRGLLELYQASLYRVH 166

QY 225 GEKVMDEAETFTSKYLREALQKIPASSILSLERDVLVEYGVWHTNLPRLERNTMDVFGQH 284

Db 167 GEDILDEAISFTTNHLSLAVASL--DHPLSEEVSHALKQSIIRGLPRVEARHVLVSV-QD 223

QY 285 TKKNAAEKLELAKLEFNIFHSIQRELBKHVGRWKD--SGSPMTFCRRHVEYYALAS 343

Db 224 IESHNKA--LLEFAKIDFNWQLHRELSEICRWKDKLDFQKLPYARDVVEGYFWIS 281

QY 344 CIAFEPQHSGLRGTQKMSHLITVLDMDYVFGTVDELELFTATIKWDPSAMECLPEYM 403

Db 282 GVTFEPQYSLGRKMLTKVIAMASIVDDYTSYATYELIPTYKAIERWDIKCIDELPEYM 341

QY 404 KGVMMVYHTVNEMARVAEKAQGRDTLNYARQAWACFDSYMOEAKWIAATGYLPTPEYL 463

Db 342 KPSYKALLDVVEEMVQLVAEHGQYRVEYAKNAMIRLAQSYLVEARWTLQNYKPSFEFK 401

QY 464 ENKVSAHRPCALPILTL-DIPFPDHLKEVDVFPKNDLICIIILRLRGDTRCYKADR 522

Db 402 ANALPTCGYAMLATISFVGMGDIVTPTFKWAANDP-KITQASTIICRFMDVVAEHKPKH 460

QY 523 ARGEASSISCYMKDNPGLTEEDALNHNFMIRDAIRLNWELLKPDNSVPITSKKHAFD 582

Db 461 RREDDCSAIECYMEY-GVTAQAYDVFNKHVESAWKDVNKEFLKP-TEMPTEVLNLSN 518

QY 583 ISRVWHHGYYRDGYSFANVETKSLVMTVIEPVL 618

Db 519 LARVMDVLYREGDGYTVVGKAAKGGITSLIIIEPTAL 554

## RESULT 4

H84633

probable limonene cyclase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001

C:Accession: H84633

R.; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: H84633  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-591 <STO>  
A:Cross-references: GB:AE002093; NID:g4115381; PIDN:AA003382.1; GSPDB:GNC0139  
C:Genetics:  
A:Gene: At2g24210  
A:Map position: 2  
C:Superfamily: vetispiradiene synthase 1

Query Match 21.2%; Score 689.5; DB 2; Length 591;  
Best Local Similarity 29.3%; Pred. No. 1.2e-40;  
Matches 182; Conservative 130; Mismatches 238; Indels 71; Gaps 19;

QY 23 ALRRTIPTGICRPGKSVAHNSIMCL-----TSVASTDSQVRVGNVYHNSLNWDDDFIOSL 77  
Db 16 ALRKTLR-----RPOSST-----CIIVTETTPCNKSPVQRRSANYQPSRWDDHHLLSV 64  
QY 78 LSTPYGADYRERADRLICEVKDIMEFNKSLDGGNDLLQRLLLVDDVERLIGDRHFKE 137  
Db 65 ENKFAKDRVRER-DLLKEKVR-----KMLNDEQKTVLDQLDFIDDLQKLGVSVHFEE 117  
QY 138 IKTALDYVNSYNEKIGCGRESVVTDLNSALGLRLTLRHGYTVSSDVNLNVFKDNGOF 197  
Db 118 IDNIL--TSSYKDR-----INIQESDLHATALEFLRFGHGFNVSEVDFVFMENCCKP 170  
QY 198 SSTANIQEIGRVLNIFRSLAVAPGKWD--BAETFTSKYLREALQKIPASSILSLE 256  
Db 171 D-----RDDIYGLISYEASLYSTKLDKNLQIFIRPFATQQLRDFVDTHSNEDEFGSD 223  
QY 257 IRDV-----LEYGWHNLRLPRLARNYVDFGQHTKKNNAEKLLELAKLEFNIFHSLORE 312  
Db 224 MVEIVQALDPYVYQMRRLSTRWIDYVGRQNYKNLV--VWEPAKIDFNIVQAIHQEE 281  
QY 313 LKHVSRRWKDSG-SPEMTFCRRHVRHVEYVALASCIAPFQHSGLRGTGMSHLITVLDDM 371  
Db 282 LKNVSSMMWETGLGQLYFARDRIVENYFWITQIQEPQYGVVQRTKINALITIDDI 341  
QY 372 YDVGTVDELELFTATIKRWDPSAMECLPEYMKGVMMVHYTHVNMARVAEKAQGRDTLN 431  
Db 342 YDIYGTLEELQLTQVAFENWDIRLDELPEYMRLLCFLVIYNEVNSIACEILRTKNINVP 401  
QY 432 YARQAEACFDSYMOEAKWIATGYLPTPEEYLENGKVSSA-----HRPCALQPLTLDI 485  
Db 402 FLKSWTDVSKAYLVEAKWYSGHKPNLEETWQARISISPTTFVHFYCVFSQLSIQV 461  
QY 486 FPDHILKEVDFPSKLDNLIC--ILRLRGDTRCYKADRARGEASSISCYMKDNPGLTE 543  
Db 462 -----LETLSQHQNVVRCSSSVFRLANDLVTSDELARGDVCKSIQCYMSET-GASE 513  
QY 544 EDALNHNFMIRDAIRELNWELLKPDNSVPTSKHAP-----DISRVWHHGYRYRGY- 597  
Db 514 DKARSHVRQMTNDLWDEMNYEKMAHSSSI-----LHDFMETVINLARMSQCMYQVGDGHG 569  
QY 598 SFANVETKSLVMRTVIEPVPL 618  
Db 570 SPEKAKIVDRVWSLNFNPL 590

RESULT 5  
A56118  
vetispiradiene synthase 1 - *Hyoscyamus muticus* (fragment)  
C:Species: *Hyoscyamus muticus*  
C:Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 29-Sep-1999  
A:Accession: A56118  
R:Back, K.; Chappell, J.  
J. Biol. Chem. 270, 7375-7381, 1995

A:Title: Cloning and bacterial expression of a sesquiterpene cyclase from *Hyoscyamus muticus*.  
A:Reference number: A56118; MUID:95221394; PMID:7706281  
A:Accession: A56118  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-520 <BAC>  
A:Cross-references: GB:U20188; NID:g763422; PIDN:AAA86337.1; PID:g763423  
C:Superfamily: vetispiradiene synthase 1

Query Match 20.6%; Score 669.5; DB 2; Length 520;  
Best Local Similarity 31.2%; Pred. No. 2.5e-39;  
Matches 161; Conservative 111; Mismatches 211; Indels 33; Gaps 13;

QY 112 GNDLQRLLLVDDVERLIGDRHFKEIKTALDYV---NSYNEKIGCGRESVWTDLNSL 168  
Db 29 GTTLTEKLNLDILIERLGIAVHFEKQIEDMLDHIYADPYF-----EAHEYNDLNTIS 80  
QY 169 ALGLRTLRLHGYTVSSDVNLNVFKDNGQSPSTANIQEIGRVLNIFRSLAVAPGK 228  
Db 81 SVQFELLRQHGYNVSPNIFSRFQDANGFKES---LRSDIRGLNLNLYEASHVRTHKEDI 136  
QY 229 MDEAETFTSKYLREALQKIPASSILSLEIRDVLEYGWHNLRLPRLARNYVDFGQHTK 288  
Db 137 LEEALVFSVGHLESAAPHL--KSPUSKQVTHALDOSLHKSIPRVEIRFISYI-EEBFFK 193  
QY 289 NAAEKLLELAKLEFNIFHSLOERELKHVSRRWKD--SGSPEMTFCRRHVRHVEYVALASCIAP 347  
Db 194 N--DLLRFAKLDYNNLLQMLHKEHSEVSRWKOLDFTVTLFYADRAVECYFTWGVYA 251  
QY 348 EPQHSGLRGTGMSHLITVLDDMYDVGTVDELELFTATIKRWDPSAMECLPEYMKGV 407  
Db 252 EPQXSQARWMLAKTIAMISIVDDTFDAYGIVKELEYVTDALQRMWDISQIDRLPEYMKISY 311  
QY 408 MMVYHTVNMARVAEKAQGRDTLNVARQAEACFDSYMOEAKWIATGYLPTPEEYLENGK 467  
Db 312 KALLDLYDDYKEKLSKGRSDVHYAKRMKEIVGNFYEGKWFTEGYMPSVSEYLSNAL 371  
QY 468 VSSAHRPCALQPLTLDIPFPDHILKEVDFPSKLI--NDLICILRLRGDTRCYKADRA 525  
Db 372 ATSTVYLLTTTSYLGMSKATKEHFEWLTATNPKELEANNATLC---RVVDDIATYVEKGRG 428  
QY 526 EASISICYKNDNGLTEEDALNHNFMIRDAIRELNWELLKPDNSVPTSK--KHAFTDI 583  
Db 429 QIATGIECYMWDY--GVSTEVAMEKFOEMADIAWKNVNEILRP---TPVSSEILTRILNL 484  
QY 584 SRVWHHGYRY--RDCGVSFANVETKSLVMRTVIEPVPL 618  
Db 485 ARIDVYKHNQDGYTHPEKVLKPHIATLVVDSIDI 520

RESULT 6  
F71434  
probable limonene cyclase - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
A:Variety: *columbia*  
C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 20-Jun-2000  
C:Accession: F71434  
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirks, P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terryn, N.; Giele, vanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B. Nature 391, 485-488, 1998  
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech, erthof, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ansc C.; Chawatzis, N.  
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*.  
A:Reference number: A71400; MUID:98121113; PMID:9461215  
A:Accession: F71434  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-600 <BEV>  
A:Cross-references: GB:Z97341; NID:g2244991; PID:g2245028  
C:Genetics:  
A:Map position: 4COP9-4G3845

RESULT 7  
T03714  
5-*epi*-aristolochene synthase - common tobacco  
N:Alternate names: sesquiterpene cyclase  
C:Species: *Nicotiana tabacum* (common tobacco)  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 29-Sep-1999  
C:Accession: T03714  
R:Facchini, P.J.; Chappell, J.  
Proc. Natl. Acad. Sci. U.S.A. 89, 11088-11092, 1992  
A:Title: Gene family for an elicitor-induced sesquiterpene cyclase in tobacco.  
A:Reference number: Z15024; MUID:93066390; PMID:1438319  
A:Accession: T03714  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-550 <PAC>  
A:Cross-references: EMBL:L04680; NID:g170342; PIDN:AAA19216.1; PID:g505588  
A:Experimental source: strain NK326  
C:Genetics: 37/3; 126/1; 251/2; 324/2; 370/3; 452/3  
A:Introns:  
C:Function:  
A:Description: mediates the conversion of the isoprenoid intermediate farnesyl  
C:Superfamily: vetispadiene synthase 1

C;Function: mediates the conversion of the isoprenoid intermediate farnesyl diphosphate to farnesol

```
Db 16 BEIRPVADSPSLWGRFLSFSDINQVETKVAQ-----ETIEPLKEQTRSMLL 63
Qy 110 DGGNDLLQRLLLVDDVERLIGDRHFKKEIKTALDYV--NSYNEKIGCGRESVVTDLNST 168
Db 64 ASGRKLSLTLNIDVIERLGIAYHFEKEIDEILDRIYENSNFEG-----DVYNEDLCTC 118
Qy 169 ALGRLTLRHGYTVSSDVLNVFKDNGQFSSTANTQIEGEIRGVNLPRASLVAFFGKGV 228
Db 119 RLOFRLRHQGYNTSLKIFSKFLDNGRLKES-----LASDVLGLLSLYEASHVRSHGEDI 174
Qy 229 MDEATSTKYLREALQKIPASSILSLERDVLVLEYGWHNLPRLEARNMVDVFGQHTKNK 288
Db 175 LEDALAFSTTHLESATPLEYP--LKEQVRHALESILHKGIPRIEIQFISVYDQKAIK 232
Qy 289 NAEKLLAKLEFNIFHLSLOERELKHVSRRWKDQSGSPE-MTFFCRRHREVEYALASCIAF 347
Db 233 N--DVLLRFKLDYNMLQMLKHQBLAEVSRWKDLNFVNTLPYARDRVCEYFALGVVY 290
Qy 348 EPQSGGRLGFTKSHLITVLDDMYDVFVGTVELELFTATIKRWDPDSAMECLPEYMGVY 407
Db 291 EPQYSQARVMLVKTIAMISIVDDTYDAYGTVDLAITYDVIQWRDIKEDISLPTYMK--- 347
Qy 408 MMVYHTVMEMARVAEKAQRD---TLNVARQAEACFDSYMOEAKWTIATGVLPTFEYVL 463
Db 348 -ISYKALLDLYKDYEKEMSRDGRSHVYVYAKERLKVSYNIEAKWFIEGHMPASEYL 406
Qy 464 ENGVSSAHRPCALQPIILTLPDPDHILKEVDFF--SKLNDLI--CI-ILRLGDRTRY 518
Db 407 RNAFTVTTYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 460
Qy 519 KADRARGEASISCYMKDNPGLTEEDALNHNFMIRDAIRELNWELLKPDNSVPISTKK 578
Db 461 EVEKNRGSLTGIECYMKDYSVSTKE-AMAKFQEMGESGWDKINRGMLRP-TPIMPFLS 518
Qy 579 HAFDISRVWHHGYRY-RDGYSFANVETKSLVMRTVIE 614
Db 519 RILNARLVVDVYKHNEDGYTHPEKVIKPHIIVVD 555

RESULT 9
T06266
germacrene C synthase - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-May-2000
C:Accession: T06266
R:Colby, S.M.; Crock, J.; Dowdle-Rizzo, B.; Lemaux, P.G.; Croteau, R.
Proc. Natl. Acad. Sci. U.S.A. 95, 2216-2221, 1998
A:Title: Germacrene C synthase from Lycopersicon esculentum cv. VFNT Cherry tomato: cDNA
A:Reference number: Z15576; MUID:98151492; PMID:9482865
A:Accession: T06266
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-548 <COL>
A:Cross-references: EMBL:AF035631; NID:g2967688; PIDN:AAC39432.1; PID:g2967689
A:Experimental source: cultivar VFNT
C:Superfamily: vetispiradiene synthase 1

Query Match 18.9%; Score 615.5; DB 2; Length 548;
Best Local Similarity 27.0%; Pred. No. 1.7e-35;
Matches 159; Conservative 136; Mismatches 229; Indels 65; Gaps 16;

Qy 52 ASTDSVQRVRGNYHNLWDDDFI-----QSLISTPYGADPYRERADRLIGEVDIMF 103
Db 3 ASSADKCRPLANFHPVSVGHHFSLVTHETITNOEKVE-----VDEYKETIKMLVETCD--- 55
Qy 104 NFKSLDGGNDLLQRLLLVDDVERLIGDRHFKKEIKTALDYV--NSYNEKIGCGRES 160
Db 56 -----NSTQKLVLDAMQRLGVAYHFHFDNEIETSIQNIQIFDASSKQNDND----- 98
Qy 161 VVTDLNSTALGLRTLRLHGYTVSSDVLNVFKDNGQFSSTANIQIEGEIRGVNLPRASL 220
Db 99 --NNLVVYSLRFLVRQOQHYSVDFKQFTNQDGKFKE-----LTNDVQGLLSLYEASH 152
Qy 221 VAFPEKVMDEAETFTSKYLREALQKIP-ASSILSLERDVLVLEYGWHNLPRLEARNYMD 279
Db 153 LRVRNEEILLEALFTTTHLESIVSNLSNNNSLKVEVEGALTPQIRMTLPRMGARKYIS 212
Qy 280 VFGQHTKKNKAAEK-LLELAKLEFNIFHLSLOERELKHVSRRWKD-SGSPMTFCRRHRE 337
Db 213 IY-----ENNDAHHLLKFKALDFNMLQKHFQRELSDLTRWKKOLDFANKYPYARDRLVE 268
Qy 338 YYALASCIAPFQSGRFLGFTKSHLITVLDDMYDVFVGTVELELFTATIKRWDPDSAME 397

Query Match 18.9%; Score 614.5; DB 2; Length 548;
Best Local Similarity 26.8%; Pred. No. 2e-35;
Matches 158; Conservative 137; Mismatches 229; Indels 65; Gaps 16;

Qy 52 ASTDSVQRVRGNYHNLWDDDFI-----QSLISTPYGADPYRERADRLIGEVDIMF 103
Db 3 ASSADKCRPLANFHPVSVGHHFSLVTHETITNOEKVE-----VDEYKETIKMLVETCD--- 55
Qy 104 NFKSLDGGNDLLQRLLLVDDVERLIGDRHFKKEIKTALDYV--NSYNEKIGCGRES 160
Db 56 -----NSTQKLVLDAMQRLGVAYHFHFDNEIETSIQNIQIFDASSKQNDND----- 98
Qy 161 VVTDLNSTALGLRTLRLHGYTVSSDVLNVFKDNGQFSSTANIQIEGEIRGVNLPRASL 220
Db 99 --NNLVVYSLRFLVRQOQHYSVDFKQFTNQDGKFKE-----LTNDVQGLLSLYEASH 152
Qy 221 VAFPEKVMDEAETFTSKYLREALQKIP-ASSILSLERDVLVLEYGWHNLPRLEARNYMD 279
Db 153 LRVRNEEILLEALFTTTHLESIVSNLSNNNSLKVEVEGALTPQIRMTLPRMGARKYIS 212
Qy 280 VFGQHTKKNKAAEK-LLELAKLEFNIFHLSLOERELKHVSRRWKD-SGSPMTFCRRHRE 337
Db 213 IY-----ENNDAHHLLKFKALDFNMLQKHFQRELSDLTRWKKOLDFANKYPYARDRLVE 268
Qy 338 YYALASCIAPFQSGRFLGFTKSHLITVLDDMYDVFVGTVELELFTATIKRWDPDSAME 397
```

## RESULT 10

T06265

germacrene C synthase, epidermal - tomato

C:Species: Lycopersicon esculentum (tomato)

C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 11-May-2000

C:Accession: T06265

R:Colby, S.M.; Crock, J.; Dowdle-Rizzo, B.; Lemaux, P.G.; Croteau, R.

Proc. Natl. Acad. Sci. U.S.A. 95, 2216-2221, 1998

A:Title: Germacrene C synthase from Lycopersicon esculentum cv. VFNT Cherry tomato: cDNA

A:Reference number: Z15576; MUID:98151492; PMID:9482865

A:Accession: T06265

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-548 &lt;COL&gt;

A:Cross-references: EMBL:AF035630; NID:g2967686; PIDN:AAC39431.1; PID:g2967687

A:Experimental source: cultivar VFNT

C:Superfamily: vetispiradiene synthase 1

Query Match 18.9%; Score 614.5; DB 2; Length 548;  
Best Local Similarity 26.8%; Pred. No. 2e-35;  
Matches 158; Conservative 137; Mismatches 229; Indels 65; Gaps 16;

Qy 52 ASTDSVQRVRGNYHNLWDDDFI-----QSLISTPYGADPYRERADRLIGEVDIMF 103  
Db 3 ASSADKCRPLANFHPVSVGHHFSLVTHETITNOEKVE-----VDEYKETIKMLVETCD--- 55  
Qy 104 NFKSLDGGNDLLQRLLLVDDVERLIGDRHFKKEIKTALDYV--NSYNEKIGCGRES 160  
Db 56 -----NSTQKLVLDAMQRLGVAYHFHFDNEIETSIQNIQIFDASSKQNDND----- 98  
Qy 161 VVTDLNSTALGLRTLRLHGYTVSSDVLNVFKDNGQFSSTANIQIEGEIRGVNLPRASL 220  
Db 99 --NNLVVYSLRFLVRQOQHYSVDFKQFTNQDGKFKE-----LTNDVQGLLSLYEASH 152  
Qy 221 VAFPEKVMDEAETFTSKYLREALQKIP-ASSILSLERDVLVLEYGWHNLPRLEARNYMD 279  
Db 153 LRVRNEEILLEALFTTTHLESIVSNLSNNNSLKVEVEGALTPQIRMTLPRMGARKYIS 212  
Qy 280 VFGQHTKKNKAAEK-LLELAKLEFNIFHLSLOERELKHVSRRWKD-SGSPMTFCRRHRE 337  
Db 213 IY-----ENNDAHHLLKFKALDFNMLQKHFQRELSDLTRWKKOLDFANKYPYARDRLVE 268  
Qy 338 YYALASCIAPFQSGRFLGFTKSHLITVLDDMYDVFVGTVELELFTATIKRWDPDSAME 397

Db 269 CYFWILGVYFEPKYSRARKMMTKVLNLTSIIDTDFDAYATFDELVTNDALQRWDANAID 328

Qy 398 CLPEYMKGVYMYVHTVNEARVAEKAQGRDITLNYARQAEACFDSYMOQAKWI-ATGYL 456

Db 329 SIQPNRPAQALDIYSEMEQVLSKGLDRVYAKNEMKULVRYFKETQWINDCDHI 388

Qy 457 PTEPYELKGVKSARHPALQALPILTDIPFPDHLKKEVDVFPKJNDLI-----CIILRL 511

Db 389 PKYEQVENAIVSAGY-----NMISTCLGVIEEFISHETFEWLMNESVIVRASALIARA 443

Qy 512 RGTCTCYKADRARGEASSICVMKONPGLTEEDALNHNF--IRDAIRELNWELLKPD 569

Db 444 MNDIVGHEDEQERGHVASLIECTMYGASKQE---TYIKFLKEVTNAWKDINKQFRP- 499

Qy 570 NSVPIITSKHAFFDISRVWHHGYRDCGYSPANVETKSLVMRTVIEPVL 618

Db 500 TEVPMFVLKVLNLRVADTLTKYKUTYTNAGKLNKMNISILIESVKI 548

RESULT 11

G71434

probable limonene cyclase - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: Columbia

C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 05-Dec-1998

C:Accession: G71434

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dick P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzengesser, T.; Pohl, T.M.; Terry, N.; Giel avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B. Nature 391, 485-488, 1998

A:Title: Sequence and analysis of chromosome 4 of Arabidopsis thaliana

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ang C.; Chludzisz, N.

A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: G71434

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1024 <BEV>

A:Cross-references: GB:Z97341; NID:g2244991; PID:e327011; PID:g2245029

C:Genetics:

A:Map position: 4COP9-4G3845

Query Match 18.0%; Score 586.5; DB 2; Length 1024;

Best Local Similarity 30.2%; Pred. No. 4.8e-33;

Matches 163; Conservative 109; Mismatches 195; Indels 73; Gaps 17;

Qy 86 DYERADRLIGEVDIMFNFKSLDGGNDLLQRLIVDDVERLIGDRHFKKEIKTALDYV 145

Db 5 DNVERTLLKQVSKML-----NETEGLEQLLEIDTLQRLGVSYHFEQEKTKLTNV 57

Qy 146 NSVWNEKIGCGRESVVT---DLNSTALGLRTLRLHGYTVSSDVLNVFKDKNGQFSSTA 201

Db 58 -----HVKNVPAKRNDRNWDGLYATALEFRLLRQH-----DV-----FDGNI 97

Qy 202 NIQIEG-EIRGVNLFRASLVAFPEKWMDEAETFTSKYLR---EALQKIPASSILSLEI 257

Db 98 GVDLDDKDIKGILSLVEASVLSRIDTKLESIVYTTKLRKFEVWKNKETSQVTLRRMV 157

Qy 258 RDVLEYGWHTNLRLEARNYMDVFGQ-HTKNKAAEKLLELAKLEFNIFHSLQERELKHV 316

Db 158 IHALEMPYHRRVGRLEARNVIEVYGERHDWN-----PILLELAKLDENFVQAIHQDELKSL 213

Qy 317 SRWKKDSG-SPEMTFCRHRHVEVYALASCIAPQHSGRFLGFTKMSHLITVLDDMDYDF 375

Db 214 SSWSKTGLTKHLDFVDRITGTEYFSSGVWYVPEFAYHROMLTKVFMILTIDDIY 273

Qy 376 GTVDELELFTATIKRWDPSAMECLPEYMKGVYMYVHTVNEARVAEKAQGRDITLNYARQ 435

Db 274 GTLEELQLFTTIVEKMDVNRLEELPNYMKLCFLCLVNEINQIGYFVFLDRKGFNVIPYLKE 333

Qy 436 AWEACFDSYMOQAKWIATGYLPTFEYLENGKVSAA-----HRPCALQPIILTDI--PF 487

Db 334 SWADMCITTELKEAKWKYKSGYKPNFBEYMQMNGWISSVPTILLHLFCLLSD-QTLDLGSY 392

Qy 488 PDHILKEVDFPKJNDLIICIILRLGD-----TRCYKADRARGEASSISCYMKD 537

Db 393 NHSVRS-----SATILRLANDLATSSVSHGFTTYNTTELARGDTMKSVQCHMHE 442

Qy 538 NPLGTEEDALNHNFIRDAIRELNWELLKPDNSVPITSKKHAFDISRVWHHGYRVDGY 597

Db 443 T-GASEAESRAYIQGIIGVADWDDLNVE--KKSCRLHQGFLEAAANLGRVAQCVCVYQGDGH 499

RESULT 12

E96723

hypothetical protein F20P5.19 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001

C:Accession: E96723

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K., ansen, N.F.; Hughes, B.; Huizar, L. Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E96723

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-632 <STO>

A:Cross-references: GB:AE005173; NID:g2194130; PIDN:AAB61105.1; GSPDB:GN00141

C:Genetics:

A:Gene: F20P5.19

A:Map position: 1

C:Superfamily: vetispiradiene synthase 1

Query Match 17.5%; Score 569; DB 2; Length 632;

Best Local Similarity 26.5%; Pred. No. 4e-32;

Matches 174; Conservative 120; Mismatches 276; Indels 86; Gaps 18;

Qy 10 VRSK-----LSSSHEIKALRRITPTLIGICRPGKVAHSINMCL-----TSVASTDSVQRRV 61

Db 16 LSLCWLRLSSSYHYPLKSLSFQSPKK-----LCIVRATNTPTDNTSTRSF 68

Qy 62 GNVHSLWDDDFQSLISTPYGADYERAD--RLIGEVKDIMFNFKSLDGGNDLLQRL 119

Db 69 THPPSLGHHFLUSASV-----NOTEMDMLWRQLEALKPIV-NAMLLPCNGADAKKIT 120

Qy 120 LLVDDVERLIGDRHFKKEIKTALDYVNSYWNEKIGCGRESVVTDLNSTALGLRTLRLHG 179

Db 121 CFHTLVLGLSVYHFEKIVEFLKDAFENIEDMIIDCKED---DLTVTSIIIFRVFLYG 176

Qy 180 YTVSS-----DVLNVFKDKNGQFSSTANIQIEGBIRGVNLNLFRA 218

Db 177 HYITPELHITISYFFFLIFHTCMVDIFNRFKGGDGNFKCLN-----DDVRGMLSFYEA 232

Qy 219 SLVAFPEKVMDEAETFTSKYLRAL--OKIPASSILSLEIRDVLEYGHTNLPLEARN 276

Db 233 SHEGTTEDLLEEMAFSTQKLELFLVGKAKHYPHITKLIQAALYIPQNFNLILVARE 292

Qy 277 YMDVFGQHTKNKAAEKLLELAKLEFNIFHSLQERELKHVSRWK--DSGSPMTFCRHR 334

Db 293 YIDFYELETDHN---EMLLKLAKLNFRFLQVLIQDLKTLITTWKELDLVSKIPVYFRER 349

Qy 335 HVEYVALASCIAPQHSGRFLGFTKMSHLITVLDDMDYFVGTVDELELFTATIKWDPS 394

Db 350 LAEPYFWATGIYEPQYSAARINLAKSIILVIDVDNFDYGTIDEVKSILVQAIERWDS 409

Qy 395 AMECLPEYMKGVYMYVHTVNEARVAEKAQGRDITLNYARQAEACFDSYMOQAKWIATG 454

Db 410 AVDVLDPYLVKVFRTFDLFKELEEYVSSEARSFTMQAYEQRLMKGLVQAEWSNRG 469



```

QY 572 VPITSKKHAFLDISRVVHHGYRDCGYSFANVETKSLVMRTVIEPVPL 618
Db 734 VVPRECKEAFKMSKVLNLFYRKDDG--FTSNDLMSLVKSVIYEPVSL 779

RESULT 15
T52059
ent-kaurene synthase (EC 2.5.1.1-) [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C:Accession: T52059
R:Yamaguchi, S.; Sun, T.; Kawaide, H.; Kamiya, Y.
Plant Physiol. 116, 1271-1278, 1998
A:Title: The GA2 locus of Arabidopsis thaliana encodes ent-kaurene synthase of gibberellin biosynthesis
A:Reference number: Z25926; MUID:98205064; PMID:9536043
A:Accession: T52059
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-785 <YAM>
A:Cross-references: EMBL:AF034774; PIDN:AAC39443.1
C:Genetics:
A:Gene: GA2
C:Keywords: transferase

Query Match 15.5%; Score 504.5; DB 2; Length 785;
Best Local Similarity 28.8%; Pred. No. 2e-27;
Matches 152; Conservative 105; Mismatches 230; Indels 41; Gaps 15;

QY 114 DLLQRLLLVDDVERLGHDFKKEIKTALDYVNSYNEKIGCGRESVVTDLNSTALGLR 173
Db 270 DQVARSIIIVTLESIGIDRDFKTEIKSILDETVRYMLR-----GDEEICLDLATCALAFR 324

QY 174 TLRHGYTVSSDVNLFVKDKNGQFSSTANTQIEGEIR--GVLNLFASLVAPEGEKVM 230
Db 325 LLLAHGYDVSYPDKLPFAESG-FSDT---LEGYKNTFSVLELFKAA-QSYPHESALK 378

QY 231 EATFTSTKYLRLEALQKIPASSI-----LSLEIRDVLEYGHTNLPRLERNYMDVFGQHTK 286
Db 379 KQCCWTQKYLEMELSWKTSVRDKYLKKEVEDALAFPSYASLERSDHRKI-LNGSAVE 437

QY 287 NKNAE-----KLELAKLEFNI FHSLOERELKHVSRWKDGSPEMTFCRHRH 335
Db 438 NTRVTKTSYRLHNICTSDILKLAVDNFCQSIHREMERLDRWIVENRLQELKFARQKL 497

QY 336 VEYVALASCIAPFQHSQPLGFTKMSHLITVLDDMDYDFGTDELELFTATIKRWDP 395
Db 498 AYCYSQAATLFGPFLSDARISWAKGCVLTIVVDVDFDVGSGKEELENLHLVEKMDLNG 557

QY 396 M-ECLPEYKGVYMMVYHTVNEARVAEKAQGRDITLNYARQAEACFDSYMQEAKWIATG 454
Db 558 VPYSSHEVELIIFSVLEDITILETGDKAFTYQGRNVTHIVKIWL DLLKSMLEAEWSSDK 617

QY 455 YLPTFEYLENGKVSAAHRPCALQILTLDPDPDHILKEVDFF--SKNDLICIILRLR 512
Db 618 STPSLEDYEMENAVISFALGPIVLPAIYLIQGPLPE---KTVDSHQYNQLYKLVSTMGRLL 674

QY 513 GDTRECKADARGEEASSISCYMK-DNPGLTEDALNHFMIRDAIRELNWELLKPDNS 571
Db 675 NDIQGFRESAEG-KUNAVSLHMKHERDKRSKEVIIESMKGLEAKRKEELHKLVLEEKGS 733

QY 572 VPITSKKHAFLDISRVVHHGYRDCGYSFANVETKSLVMRTVIEPVPL 618
Db 734 VVPRECKEAFKMSKVLNLFYRKDDG--FTSNDLMSLVKSVIYEPVSL 779

```

Search completed: July 7, 2004, 17:44:16  
Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:38:32 ; Search time 17 Seconds  
(without alignments)  
1892.903 Million cell updates/sec

Title: US-10-025-145A-65

Perfect score: 3251

Sequence: 1 MALLSITPLVRSRSLSSSHE.....FANVETKSLVMRTVIEPVPL 618

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2453	75.5	628	1 TSD1_ABIGR	O24475 abies grand
2	2200.5	67.7	627	1 TSD2_ABIGR	O24474 abies grand
3	2031.5	62.5	637	1 TSD3_ABIGR	O22340 abies grand
4	1240.5	38.2	862	1 TASY_TAXBR	Q41594 taxus brevi
5	1238.5	38.1	862	1 TASY_TAXBA	Q93ya3 taxus bacca
6	1209.5	37.2	862	1 TASY_TAXCH	Q9ft37 taxus chine
7	717.5	22.1	554	1 DCS1_GOSAR	Q39761 gossypium a
8	717.5	22.1	554	1 DCS2_GOSAR	Q39760 gossypium a
9	701.5	21.6	554	1 DCS4_GOSAR	O49853 gossypium a
10	699.5	21.5	555	1 DCS3_GOSAR	Q43714 gossypium a
11	692.5	21.3	554	1 DCS1_GOSHI	P93665 gossypium h
12	667	20.5	601	1 CASS_RICCO	P59287 ricinus com
13	644.5	19.8	548	1 SEAS_TOBAC	O40577 nicotiana t
14	133	4.1	1941	1 UBR1_KLULA	O60014 kluyveromyc
15	125	3.8	1295	1 BXA2_CLOBO	Q45894 clostridium
16	117.5	3.6	1513	1 GRUF_RAT	P81128 rattus norv
17	114	3.5	456	1 LPP8_ECOLI	P25739 escherichia
18	111.5	3.4	602	1 LPPA_HELPY	O25122 helicobacte
19	110	3.4	2875	1 RPLP_TSWVL	P28976 tomato spot
20	109.5	3.4	1513	1 GRLF_HUMAN	Q9ary4 homo sapien
21	108.5	3.3	570	1 EGRA_BACAA	O81kt8 bacillus ap
22	108.5	3.3	1061	1 EX5C_BUCAP	Q8K9b0 buchnera ap
23	107	3.3	815	1 CS53_YEAST	Q12018 saccharomyc
24	105.5	3.2	428	1 SVS_PASMU	P57836 pasteurella
25	105.5	3.2	898	1 YMV6_YEAST	Q04748 saccharomyc
26	105.5	3.2	1026	1 MY1B_DROME	O23979 drosophila
27	105	3.2	1790	1 VIT_ANTGR	Q05808 anthonomus
28	104.5	3.2	510	1 CHLB_PINTH	Q00864 pinus thunb
29	104.5	3.2	570	1 EZFA_BACCR	O817a9 bacillus ce
30	104.5	3.2	1500	1 GRKF_CANFA	P83509 canis famil
31	104	3.2	1510	1 MUKB_HAEIN	P45187 haemophilus
32	103.5	3.2	598	1 SYD_LACPL	Q88vq8 lactobacill
33	103.5	3.2	822	1 V2B_BMW	P03594 brome mosai

#### RESULT 1

ID	TSD1_ABIGR	STANDARD;	PRT;	628 AA.
AC	O24475;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Pine synthase, chloroplast precursor (EC 4.2.3.14) (Beta-geraniolene synthase) ((-)-(1S,5S)-pinene synthase).			
GN	AG3.18.			
OS	Abies grandis (Grand fir).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.			
OX	NCBI_TaxID=46611;			
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RX	MEDLINE=97413772; PubMed=9268308;			
RA	Bohlmann J., Steele C.L., Croteau R.;			
RT	"Monoterpene synthases from grand fir (Abies grandis). cDNA isolation, characterization, and functional expression of myrcene synthase, (-)-(4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase.";			
RL	J. Biol. Chem. 272:21784-21792(1997).			
CC	FUNCTION: Involved in defensive oleoresin formation in conifers in response to insect attack or other injury. Involved in monoterpene (C10) olefins biosynthesis. A mixture of alpha-and beta-pinene is produced by this enzyme.			
CC	COFACTOR: Manganese and potassium.			
CC	COFACTOR: Oleoresinosis.			
CC	SUBCELLULAR LOCATION: Chloroplast.			
CC	INDUCTION: By wounding.			
CC	SIMILARITY: Belongs to the terpene synthase family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; U87909; AAB71085.1; -.			
DR	HSSP; Q40577; SEAT.			
DR	InterPro; IPR008930; Terp_cyc_toroid.			
DR	InterPro; IPR001906; Terp_synth-like.			
DR	InterPro; IPR005630; Terpene synth C.			
DR	InterPro; IPR008949; Terpenoid synth.			
DR	Pfam; PF01397; Terpene synth; 1.			
DR	Pfam; PF03936; Terpene synth_C; 1.			
KW	Lyase; Manganese; transit peptide; Chloroplast.			
FT	TRANSIT 1 ?			
FT	CHAIN 1 ?			
FT	ACT_SITE 524 524			
FT	ACT_SITE 601 601			
FT	ACT_SITE 605 605			
FT	SEQUENCE 628 AA; 71505 MW; 23DBB788F3C8072C CRC64;			

#### ALIGNMENTS

34	103	3.2	6669	1	NEBU_HUMAN	P20929 homo sapien
35	102.5	3.2	567	1	YE28_METUA	Q58823 methanococc
36	102.5	3.2	1085	1	CASO_BOVIN	P35384 bos taurus
37	102.5	3.2	2364	1	SPCR_HUMAN	Q01082 homo sapien
38	102	3.1	592	1	SYD_THETN	Q8rai7 thermoanaer
39	102	3.1	700	1	LSHR_RAT	P16235 rattus norv
40	102	3.1	4036	1	RRPL_DUGBV	Q66431 dugbe virus
41	101.5	3.1	692	1	IF2_OCEIH	Q86eq1 oceanobacil
42	101.5	3.1	820	1	SUS2_TULGE	Q41607 tulipa gesn
43	101	3.1	1277	1	IF3X_YEAST	Q03690 saccharomyc
44	100.5	3.1	418	1	YB87_SCHPO	O14303 schizosacch
45	100.5	3.1	519	1	YMP4_YEAST	Q04347 saccharomyc

```

Query Match      75.5%; Score 2453; DB 1; Length 628;
Best Local Similarity 75.3%; Pred. No. 3e-158;
Matches 475; Conservative 59; Mismatches 81; Indels 16; Gaps 6;

QY 1 MALLSTIPVRSCL-----SSSHEIKALRRTIPTLIGICRPGKSVHAHNMCLTSTVASTD 55
DB 1 MALVSTAPLASKCLKLSLTSSTHEIKALSRKTIPTALGMSRRGKSIPTSISSMTTIVTDD 60

QY 56 SVORRNVGNHSLWDDDFQSLISTPYGAPDYERADRLIGEVKDMFNFKSLDGG--- 112
DB 61 GVRRRMGDFHSLWDDDFQSL-PTAYEKSYLEAEKLGIVKN-MFNSMSLEDGELMS 118

QY 113 --NDLLQRLVDDVERLIGIDRHFKEIKTALDYVNSYMKKEGICGREGSVWTDLNSTAL 170
DB 119 PLNDLQRLVDDVERLIGIDRHFKEIKTALDYVNSYMKKEGICGREGSVWTDLNSTAL 178

QY 171 GLRTLRHGYTVSSDVLNVFKKNGQFSTANTQIEGEIRGVNLFRASLVAFPGEKVM 230
DB 179 GLRTLRHGYTVSSDVLNVFKKNGQFSTANTQIEGEIRGVNLFRASLVAFPGEKIMD 238

QY 231 EAEFTSKYLREALQKIPASSIISLEIRDLVLEYGWHTNLPRLRARNYVFGQHTKNA 290
DB 239 EAEFTSKYLREALQKIPASSIISLEIRDLVLEYGWHTNLPRLRARNYVFGQHTENTKS 297

QY 291 ---AEKLELLEKLEFNI FHSLOERLKHVSRRWKDGSPEMTFCRRHVEYVALASCI 347
DB 298 YVSKKLELLEKLEFNI FHSLOERLKHVSRRWKDGSPEMTFCRRHVEYVALASCI 357

QY 348 EPOHSGFRIGFTKMSHLITVLDMDYVFGTVDBLELFTATIKWDPSAMECLPEYMKGVY 407
DB 358 EPOHSGFRIGFTKMSHLITVLDMDYVFGTVDBLELFTATIKWDPSAMECLPEYMKGVY 417

QY 408 MMVYHTVNMARVAEKAQGRDNLNARQAEACFDSYMOEAKVIATGYLPTPEYLENGK 467
DB 418 IAVDVTNEMARVAEKAQGRDNLNARQAEACFDSYMOEAKVIATGYLPTPEYLENGK 477

QY 468 VSSAHRPCALQPIITLIDIPPPDHILKEVDPPSKINDLICIILRLGRDTRCYKADRAGEE 527
DB 478 VSCGHRISALQPIITLIDIPPPDHILKEVDPPSKINDLICALILRLGRDTRCYKADRAGEE 537

QY 528 ASSISCYMKNDPGLTEEDALNHNFMIRDAIRELNWELLKPDNSVPITSKKHAFDISRVW 587
DB 538 ASSISCYMKNDPGLTEEDALNHNFMIRDAIRELNWELLKPDNSVPITSKKHAFDIARAF 597

QY 588 HGVYRDGYSEANVETKSLVMTVTEPVL 618
DB 598 HGVYRDGYSEANVETKSLVMTVTEPVL 628

RESULT 2
TSD2 ABIGR STANDARD; PRT; 627 AA.
ID TSD2 ABIGR STANDARD; PRT; 627 AA.
AC Q2474;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myrcene synthase, chloroplast precursor (EC 4.2.3.15).
GN AG2-2.
OS Abies grandis (Grand fir).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=46611;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=97413772; PubMed=9268308;
RA Bohlmann J., Steele C.L., Croteau R.;
RT "Monoterpene synthases from grand fir (Abies grandis). cDNA isolation,
RT characterization, and functional expression of myrcene synthase, (-)-
RT (4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase."
RJ J. Biol. Chem. 272:21784-21792 (1997).
RC CC
CC -!- FUNCTION: Involved in defensive oleoresin formation in conifers in
CC response to insect attack or other injury. Involved in monoterpene

```

```

(C10) olefins biosynthesis.
-!- CATALYTIC ACTIVITY: Geranyl diphosphate = myrcene + diphosphate.
-!- COFACTOR: Manganese and potassium.
-!- PATHWAY: Oleoresinosis.
-!- SUBCELLULAR LOCATION: Chloroplast.
-!- INDUCTION: By wounding.
-!- SIMILARITY: Belongs to the terpene synthase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL; U87908; AAB71084.1; -.
CC HSSP; Q40577; SEAS.
CC InterPro; IPR008930; Terp cyc toroid.
CC InterPro; IPR001906; Terp synth-like.
CC InterPro; IPR005630; Terpene synth C.
CC InterPro; IPR008949; Terpenoid synth.
CC Pfam; PF01397; Terpene synth; 1.
CC Pfam; PF03936; Terpene synth C; 1.
CC Lyase; Manganese; Transist peptide; Chloroplast.
CC TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
CC CHAIN ? 627 MYRCENE SYNTHASE.
CC FT ACT_SITE 523 523 BY SIMILARITY.
CC FT ACT_SITE 600 600 BY SIMILARITY.
CC FT ACT_SITE 604 604 BY SIMILARITY.
CC SQ SEQUENCE 627 AA; 72478 MW; 2E0DA492E0C971FD CRC64;

Query Match      67.1%; Score 2200.5; DB 1; Length 627;
Best Local Similarity 68.1%; Pred. No. 3.4e-141;
Matches 427; Conservative 77; Mismatches 106; Indels 17; Gaps 7;

QY 1 MALLSTIPVRSCL-----SSSHEIKALRRTIPTLIGICRPGKSVHAHNMCLTSTVASTD 55
DB 1 MALVSTAPLASKCLKLSLTSSTHEIKPVPYRTIPTNLGMRGRGKSVTPSMSISLATAAPDD 60

QY 56 SVORRNVGNHSLWDDDFQSLISTPYGAPDYERADRLIGEVKDMFNFKSLDGG--- 112
DB 61 GVRRIIDGSHSNWDDDFQSL-STPYGEPYQERAEALIVEVKTI-FNSMYLDDORLMS 118

QY 113 --NDLLQRLVDDVERLIGIDRHFKEIKTALDYVNSYMKKEGICGREGSVWTDLNSTAL 170
DB 119 SFNDLMQRLVDDVERLIGIDRHFKEIKTALDYVNSYMKKEGICGREGSVWTDLNSTAL 178

QY 171 GLRTLRHGYTVSSDVLNVFKKNGQFSTANTQIEGEIRGVNLFRASLVAFPGEKVM 230
DB 179 GPRTLRHLGYTVSPVLKAFQDQNGQFVCSGP-QTEGEIRSVNLNFRASLVAFPGEKVM 237

QY 231 EAEFTSKYLREALQKIPASSIISLEIRDLVLEYGWHTNLPRLRARNYVFGQHTK---N 287
DB 238 EAEFTSKYLREALQKIPASSIISLEIRDLVLEYGWHTNLPRLRARNYVFGQHTK---N 296

QY 288 KNAEKLLEKLEFNI FHSLOERLKHVSRRWKDGSPEMTFCRRHVEYVALASCI 347
DB 297 KNAGKLELLEKLEFNI FHSLOERLKHVSRRWKDGSPEMTFCRRHVEYVALASCI 356

QY 348 EPOHSGFRIGFTKMSHLITVLDMDYVFGTVDBLELFTATIKWDPSAMECLPEYMKGVY 407
DB 357 DPKHSAFRLGFAKMCCHLVTLDDIYDTFGTIDLELFTSAIKRWSSEIHLPEYMKGVY 416

QY 408 MMVYHTVNMARVAEKAQGRDNLNARQAEACFDSYMOEAKVIATGYLPTPEYLENGK 467
DB 417 MVVFVTVNEMARVAEKAQGRDNLNARQAEACFDSYMOEAKVIATGYLPTPEYLENGK 476

QY 468 VSSAHRPCALQPIITLIDIPPPDHILKEVDPPSKINDLICIILRLGRDTRCYKADRAGEE 527
DB 477 VSSAYRVATLQPIITLIDIPPPDHILKEVDPPSKINDLICIILRLGRDTRCYKADRAGEE 536

QY 528 ASSISCYMKNDPGLTEEDALNHNFMIRDAIRELNWELLKPDNSVPITSKKHAFDISRVW 587

```

```
Db 537 ASCSCKMONGPSTEDDALHINAMVNDIIKELNELLRSNDNIPMLAKKHAEDITRAL 596
Qy 588 HHGYRDRGYSFANVETKSLVMRTVIE 614
Db 597 HHLYYRDRGYSFANVETKSLVMRTVIE 623

RESULT 3
TS3 ABIGR STANDARD; PRT; 637 AA.
AC O22340;
DT 28-FEB-2003 (Rel. 41, Created)
DT 10-OCT-2003 (Rel. 41, Last sequence update)
DE (-)-(4S)-limonene synthase, chloroplast precursor (EC 4.2.3.16).
GN AG10.
OS Abies grandis (Grand fir).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=46611;
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=97411372; PubMed=9268308;
RA Bohlmann J., Steele C.L., Croteau R.;
RT "Monoterpene synthases from grand fir (Abies grandis). cDNA isolation,
RT characterization, and functional expression of myrcene synthase, (-)-(4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase.";
RL J. Biol. Chem. 272:21784-21792(1997).
CC -!- FUNCTION: Involved in defensive oleoresin formation in conifers in
CC response to insect attack or other injury. Involved in monoterpene
CC (C10) olefins biosynthesis.
CC -!- CATALYTIC ACTIVITY: Geranyl diphosphate = limonene + diphosphate.
CC -!- COFACTOR: Manganese and potassium.
CC -!- PATHWAY: Oleoresinosis.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- INDUCTION: By wounding.
CC -!- SIMILARITY: Belongs to the terpene synthase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).

EMBL; AF006193; AAB70907.1; -.
DR InterPro; IPR008930; Terp_cyc_toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR InterPro; IPR005630; Terpene_synth_C.
DR InterPro; IPR008949; Terpenoid_synth.
DR Pfam; PF01397; Terpene_synth; 1.
DR Pfam; PF03936; Terpene_synth; C.
DR Lysase; Manganese, Transit peptide; Chloroplast.
FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
FT CHAIN 1 637 (-)-(4S)-LIMONENE SYNTHASE.
FT ACT_SITE 533 533 BY SIMILARITY.
FT ACT_SITE 610 610 BY SIMILARITY.
FT ACT_SITE 614 614 BY SIMILARITY.
SQ SEQUENCE 637 AA; 73477 MW; 8E80CD9DDE886898 CRC64;

Query Match 62.5%; Score 2031.5; DB 1; Length 637;
Best Local Similarity 63.6%; Pred. No. 9.2e-130;
Matches 409; Conservative 84; Mismatches 119; Indels 31; Gaps 14;

Qy 1 MALLSITPL-VSRSC-----LSSSHEIKAL--RRIPITGLICRPGKSVASHSINMCLTSA 52
Db 1 MALLSIVSLQPKSGCLKSLSSSNVQKALCLISTAVPTLRMRROKALV--INMKLTATVS 58
Qy 53 STDS-----VORRVGNHNSLWDDFIQSLISTPYGAPDYRERADRLGEVKDIMEFNE-- 105
Db 59 HRDNGGVQLQRIADHFNHNLWEDDFIQSL--SSPYGGSSYSERAEVTVEEVEKE-MFNISP 116
```

```
Qy 106 --KSELDGNDLLQRLLLVDDVERLIGDRHFKEIKTALDYNSYNNK--GIGCGRESV 162
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 NNRELFGSQNDLLRLWMVDSIERLIGDRHFQNEIRVALDVVYSYKKEGICGGRDSTF 176
Qy 163 TDINSTALGLRTIRLHGTYVSSDVLPVFKDKNQGSSTANTQIEGEI--RGVNLNFRASLV 221
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
177 PDINSTALALFTLRILHGYNVSSDVLEVPFKDKGHFACPA--ILTEGQITRSLNLYRASLV 235
Qy 222 APGEKVMDEATEFTSKYLREALOKIPASSLSLSEIRDLVLEYGWHNTNLPLEARNYMDVF 281
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
236 APGEKVMDEAEIIFSASVYKVLQKIPVSN--LSGEIEVLEYGWHNTNLPLEARNYIEVY 294
Qy 282 ---GQHTKNNA---AEKLELAKLEFNIFHSLOERELKHVSRWVKSGSGSEMTFCRRH 335
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
295 EOGVESLNEPYNMNMKLLQAKLEFNIFHSLOERELQSLSRWVKSGSSQLFTFRHRH 354
Qy 336 VEYVALASCIAPFEPQHSGLFTKMSHLIVLDMDYVFGTVDELELFTATIKWDPSA 395
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
355 VEYTMASCI SMLPKHSAPRMEFVKVCHLVTLVDIYDTFGTMNELQLFTDAIKRWDLST 414
Qy 396 MECLPEYMKGVMMVYHTVNEMARVAEKAQGRDITLNARQAEACFDSYMOEAKWIATGY 455
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
415 TRWLPPEYMKGVMDLYQCINEMVEAEKQGRDMLNLYIQNAWEALFTFMQEAKEWISSY 474
Qy 456 LPTFEYVLENGKVSASRHPICALQILTLDIPPDHILKEVDPPSKINDLICIILRLRGT 515
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
475 LPTFEYVLENGKVSASRHPICALQILTLDIPPDHILKEVDPPSKINDLICIILRLRGT 534
Qy 516 RCYKADRRARGEASSISCYMKDNLTEEDALNINFMIRDAIRLWELLKPDNSVIT 575
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
535 RCYKADRRARGEASSISCYMKDHPGSIEDALNINAMISDAIRLWELLRDPDSKSPIS 594
Qy 576 SKKHAFDLSRVHGYRDRGYSFANVETKSLVMRTVIEPVPL 618
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
595 SKKHAFDITRAFHVYKRDGYTYSNNETKLVMTKLVLEPLAL 637
```

## RESULT 4

```
TASY TAXBR
ID TASY TAXBR STANDARD; PRT; 862 AA.
AC Q41594; Q94FV8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Taxadiene synthase (EC 4.2.3.17) (Taxa-4(5),11(12)-diene synthase).
GN TDC1.
OS Taxus brevifolia (Pacific yew).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.
OX NCBI_TaxID=46220;
RN SEQUENCE FROM N.A.
RX MEDLINE=96119163; PubMed=8621577;
RA Wildung M.R., Croteau R.B.;
RT "A cDNA clone for taxadiene synthase, the diterpene cyclase that
RT catalyzes the committed step of taxol biosynthesis.";
RL J. Biol. Chem. 271:9201-9204(1996).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21297238; PubMed=11404343;
RA Trapp S.C., Croteau R.B.;
RT "Genomic organization of plant terpene synthases and molecular
RT evolutionary implications.";
RL Genetics 158:811-832(2001).
CC -!- FUNCTION: Catalyzes the cyclization of the ubiquitous isoprenoid
CC intermediate geranylgeranyl diphosphate to taxa-4,11-diene, the
CC parent olefin with a taxane skeleton.
CC -!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = taxa-4,11-diene +
CC diphosphate.
CC -!- PATHWAY: Taxol biosynthesis; first step.
CC -!- SIMILARITY: Belongs to the terpene synthase family.
```



Db 519 DDDVWQKTYLRMPSPSLNSKCLAKLDFNVQSLHQBELKLLTRWWKSGMADINETR 578  
Qy 333 HRHVEYVALASCIAPFPOHSGFRIGFTQKWSHLITVLDMDYDVGTVDELELFTATIKWD 392  
Db 579 HRVAEYV--FSSATFEPEYSATRIAFKIGCLOVLFDMDADIFATLDELKSFTEGVKWD 636  
Qy 393 PSAMECLPEYKVMYVHTVNMARVAEKAQGRDITLNYARQAWAECFDSYMOBAKWA 452  
Db 637 TSLHIEIPECMQTCFKWFKLMEEVNDVVKVQGRDMLAHIRKPWELYNFCYVQREWLE 696  
Qy 453 TGYLPTFEYLENGKVSAAHPCALQPLITLIDPPDHILKEVDPPSKLNDLICILRLR 512  
Db 697 AGYIPTFEYLUKTVSAISVGLGECTLOPILLMGLVKDDVVEKVVHPSNMFLVLSWELT 756  
Qy 513 GDTTCYKADRAARGEASSICVMKDNPGLTBEDALNHNFMIRDAIRELNWELLPKDNV 572  
Db 757 NDTKTYQAEKARGQQAAGIACVMKDNPGATEDEAKHICRVVDRLAKASFEYFKPSNDI 816  
Qy 573 PITSKHAFDISRVHHGYRDRDGSFANVETKSLVMTVIEPV 616  
Db 817 PMGCKSFIFNLRLCQVIFYKFDIGYIANEEIKDVKIRKVIDPI 860

## RESULT 6

TASY\_TAXCH STANDARD; PRT; 862 AA.  
AC Q9FT37;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Taxadiene synthase (EC 4.2.3.17) (Taxa-4 (5),11 (12)-diene synthase).  
GN TDC1.  
OS Taxus chinensis (Chinese yew).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.  
OX NCBI\_TaxID=29808;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Callus;  
RA Wang W., Shi Q., Ouyang T., Zhu P., Cheng K.;  
RT "Cloning, expression, and characterization of taxadiene synthase, a  
diterpene cyclase from Taxus chinensis";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Catalyzes the cyclization of the ubiquitous isoprenoid  
intermediate geranylgeranyl diphosphate to taxadiene, the  
parent olefin with a taxane skeleton.  
CC -!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = taxa-4,11-diene +  
diphosphate.  
CC -!- PATHWAY: Taxol biosynthesis; first step.  
CC -!- SIMILARITY: Belongs to the terpene synthase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AY007207; AAC02257.1; --  
HSPB; Q40577; SEAU.  
DR InterPro: IPR008930; Terp\_cyc\_toroid.  
DR InterPro: IPR001906; Terp\_synth-like.  
DR InterPro: IPR005630; Terpene synth C.  
DR InterPro: IPR008949; Terpenoid synth.  
DR Pfam: PF01397; Terpene synth; 1.  
DR Pfam: PF03936; Terpene synth\_C; 1.  
KW Taxol biosynthesis; Lyase.  
FT ACT SITE 758 758 BY SIMILARITY.  
FT ACT SITE 835 835 BY SIMILARITY.  
FT ACT SITE 839 839 BY SIMILARITY.  
SQ SEQUENCE 862 AA; 98069 MW; 3A597BAF722BF679 CRC64;

Query Match 37.2%; Score 1209.5; DB 1; Length 862;  
Best Local Similarity 41.9%; Pred. No. 4.9e-74;  
Matches 270; Conservative 110; Mismatches 209; Indels 55; Gaps 12;  
Qy 15 LSSGHEI--KALRRTIPTLIGICRPG-----KSAVHSINMCLTSV-ASTDSVQRVGVNYS 66  
Db 230 LSPDFEILPALLQAKALGINLPYDLPFIKYSTTREATLTDVSAADINPANNMNALE 289  
Qy 67 NL-----W-----DDDFIQSLISTPYGAPDYRERADRLIGEVKDIMFNFKS--LE 109  
Db 290 GLEFVMDWKIMRFSQKSGSFLSSPAST-----ACVLMNTGDEKCTFLNLLVVK 339  
Qy 110 DGG-----NDLLQRLLLVDVERLGDHFKKEIKETALDYVNSYWNKEGIGCGRESV 162  
Db 340 FGGCVPCMYSIDLLERLSLVNIEHLGIGRHFQKQEIKVALDVYVRHWSERGIGWRDILV 399  
Qy 163 TDLNSTALGLTTLHLGYTVSSDVLNVEKDKNGQFSSTANIQIEGIRGLNLRASLVA 222  
Db 400 PDLNNTALGLRTLRTGHGYDVSSDVLNFKDENGRAFFSAG--QTHVELRSVLLFRASDLA 458  
Qy 223 FPGKVMDEAETFTSKYLREAL--QKIPASSILSLERDVLVEYGMHTNLPRLRNMYDMVF 281  
Db 459 FPEGAMDARKEAPYLRDALATKISTNTKLFKEIEYVVEYVPHMSIPRSEARSVIDSY 518  
Qy 282 GQ-----HTKNNAEKLLELAKLEFNI FHSLOERELKHVSRWKDSGSPMTFCR 332  
Db 519 DDDYVWERKTLYRMPSPSLNSKCLAKLDFNVQSLHQBELKLLTRWWKSGMADINETR 578  
Qy 333 HRHVEYVALASCIAPFPOHSGFRIGFTQKWSHLITVLDMDYDVGTVDELELFTATIKWD 392  
Db 579 HRVAEYV--FSSATFEPEYSATRIAFKIGCLOVLFDMDADIFATLDELKSFTEGVKWD 636  
Qy 393 PSAMECLPEYKVMYVHTVNMARVAEKAQGRDITLNYARQAWAECFDSYMOBAKWA 452  
Db 637 TSLHIEIPECMQTCFKWFKLMEEVNDVVKVQGRDMLAHIRKPWELYNFCYVQREWLD 696  
Qy 453 TGYLPTFEYLENGKVSAAHPCALQPLITLIDPPDHILKEVDPPSKLNDLICILRLR 512  
Db 697 AGYIPTFEYLUKTVSAISVGLGECTLOPILLMGLVKDDVVEKVVHPSNMFLVLSWELT 756  
Qy 513 GDTTCYKADRAARGEASSICVMKDNPGLTBEDALNHNFMIRDAIRELNWELLPKDNV 572  
Db 757 NDTKTYQAEKARGQQAAGIACVMKDNPGATEDEAKHICRVVDRLAKASFEYFKPSNDI 816  
Qy 573 PITSKHAFDISRVHHGYRDRDGSFANVETKSLVMTVIEPV 616  
Db 817 PMGCKSFIFNLRLCQVIFYKFDIGYIANEEIKDVKIRKVIDPI 860

## RESULT 7

DCS1\_GOSAR STANDARD; PRT; 554 AA.  
AC Q39761;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE (+)-delta-cadinene synthase isozyme XC1 (EC 4.2.3.13) (D-cadinene  
synthase).  
OS Gossypium arboreum (Tree cotton).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eucotids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
OX NCBI\_TaxID=29729;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nanking;  
RX MEDLINE=96132653; PubMed=8554317;  
RA Chen X.-Y., Chen Y., Heinsteins P., Davison V.J.;  
RT "Cloning, expression, and characterization of (+)-delta-cadinene  
synthase: a catalyst for cotton phytoalexin biosynthesis.";  
RL Arch. Biochem. Biophys. 324:255-266(1995).  
CC -!- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl  
diphosphate (fpp) to (+)-delta cadinene.

```

CC -!- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-
CC delta-cadinene + diphosphate.
CC -!- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
CC first (committed) step.
CC -!- SIMILARITY: Belongs to the terpene synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC -----
CC EMBL: U23206; AAA93064.1; -.
CC PIR: S68365; S68365.
CC HSSP: Q40577; SEAU.
CC InterPro: IPR008930; Terp cyc toroid.
CC InterPro: IPR001906; Terp synth-like.
CC InterPro: IPR005630; Terpene synth C.
CC InterPro: IPR008949; Terpenoid synth.
CC Pfam: PF01397; Terpene synth; 1.
CC Pfam: PF03936; Terpene synth; 1.
CC Lyase; Multigene family.
CC ACT_SITE 451 BY SIMILARITY.
CC ACT_SITE 527 BY SIMILARITY.
CC ACT_SITE 531 BY SIMILARITY.
CC SEQUENCE 554 AA; 64137 MW; 59D6922DED9DCAF CRC64;

Query Match 22.1%; Score 717.5; DB 1; Length 554;
Best Local Similarity 31.9%; Pred. No. 4.7e-41;
Matches 183; Conservative 107; Mismatches 243; Indels 43; Gaps 14;

QY 51 VASTDSQVRVGNHYNLWDDDFIQSLISTPYCAPD-----YREADRLLIGVKIMFN 104
Db 14 LSNKDEMPKADQPSIWGDLFL-----NCPDKNIDAEFKRHQQLKEVR----- 60
QY 105 FKSLEDGNDLLQRLLLVDDVERLIGDRHFKKEIKTALDVNSYWNKEGICGCRSVVTD 164
Db 61 -KMWVAPMANSTQKLAFLDSVQELGVSYHFTKEIDELE--NIYHN-----NDAEND 110
QY 165 LNSTALGLRTLRLHGVTVSSDVNLNFKDNKGQSSSTANTQIEGEIRGVMLFRASLVAPP 224
Db 111 LYTTSIRFRLLRHGHVNSCDVFNKFKQGNFKSS-----VTSVDRGLLELYQASVLRVH 166
QY 225 GEKVMDEATFSTKYLREALQKIPASSILSLERDVLVYGWHTNLPRLEARNYMDVFGQH 284
Db 167 GEDILDEALSTFTHHLSLAVASL--DHPUSEVSHALKQSIIRGLPRVEARHYSVY-QD 223
QY 285 TKKNAAEKLELAKLEFNIFHSLOBERLKHVGRWKD--SGSPMTFCRRHHVEYYALAS 343
Db 224 IESHKA--LLEFAKIDFNMLQFLHRKELSEICRWKDLDFQKLPYARDVVVEGYFWIS 281
QY 344 CTAFEPQHSFRLGFTGMGHLITVLDDMDVDFGTVDLELFTATIKRWDPSAMECPYEM 403
Db 282 GVYFEPQYSLGRKMLTKVIMASISVDVTSYATYEEILPYTNAIRWDIKCIDEIPEYM 341
QY 404 KGYVMVYHTVNMARVAEKAQGRDTLNVARQAEACFDSYMOEAKWIATGYLPTFEYL 463
Db 342 KESYKALLDVEEMVQLVAEHGQYVEYAKNAMIRLAQSYLVAEKWTQNYKPSFEFK 401
QY 464 ENGVSAHRPCALQILTL-DIPFPDHLIKKEYDFPSKLNLDLCIILRLGDRTRCYKADR 522
Db 402 ANALPTCGVAMLAITSFVGMDIVTPTFKWAASDP-KIIQASTIICRFMDVVAEHKFKH 460
QY 523 ARGEERASSICYNKDPGLTEEDALNHNPMIRDAIRLWELLKPDNSVPITSKGHAFD 582
Db 461 RREDDCSAIECYMEBY-GVTAQAYDVFNKHVESAWKDLNQEFLEK-TEMPTEVLNRSLN 518
QY 583 ISRVWHGVRDGYSPANYETKSLVMRTVIEPVPL 618
Db 519 LARVMDVLYREGDGYTVGKAAGGITSLLIEPIAL 554

```

```

RESULT 8
DCS2 GOSAR
ID DCS2_GOSAR STANDARD; PRT; 554 AA.
AC Q39760;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE (+)-delta-cadinene synthase isozyme XCL4 (EC 4.2.3.13) (D-cadinene
DE synthase).
OS Gossypium arboreum (Tree cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Malvales; Malvaceae; Gossypium.
OX NCBI_TaxID=29729;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nanking;
RX MEDLINE=96132653; PubMed=8554317;
RA Chen X.-Y., Chen Y., Heinstein P., Davisson V.J.;
RT "Cloning, expression, and characterization of (+)-delta-cadinene
RT synthase: a catalyst for cotton phytoalexin biosynthesis.";
RL Arch. Biochem. Biophys. 324:255-266(1995).
CC -!- FUNCTION: Responsible for the cyclization of trans-farnesyl
CC diphosphate (FPP) to (+)-delta cadinene.
CC -!- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-
CC delta-cadinene + diphosphate.
CC -!- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
CC first (committed) step.
CC -!- SIMILARITY: Belongs to the terpene synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC -----
CC EMBL: U23205; AAA93065.1; -.
CC PIR: S68366; S68366.
CC HSSP: Q40577; SEAU.
CC InterPro: IPR008930; Terp cyc toroid.
CC InterPro: IPR001906; Terp synth-like.
CC InterPro: IPR005630; Terpene synth C.
CC InterPro: IPR008949; Terpenoid synth.
CC Pfam: PF01397; Terpene synth; 1.
CC Pfam: PF03936; Terpene synth; 1.
CC Lyase; Multigene family.
CC ACT_SITE 451 BY SIMILARITY.
CC ACT_SITE 527 BY SIMILARITY.
CC ACT_SITE 531 BY SIMILARITY.
CC SEQUENCE 554 AA; 64158 MW; A88974665E0F6B2B CRC64;

Query Match 22.1%; Score 717.5; DB 1; Length 554;
Best Local Similarity 31.9%; Pred. No. 4.7e-41;
Matches 184; Conservative 108; Mismatches 241; Indels 43; Gaps 14;

QY 51 VASTDSQVRVGNHYNLWDDDFIQSLISTPYCAPD-----YREADRLLIGVKIMFN 104
Db 14 LSNKDEMPKADQPSIWGDLFL-----NCPDKNIDAEFKRHQQLKEVR----- 60
QY 105 FKSLEDGNDLLQRLLLVDDVERLIGDRHFKKEIKTALDVNSYWNKEGICGCRSVVTD 164
Db 61 -KMWVAPMANSTQKLAFLDSVQELGVSYHFTKEIDELE--NIYHN-----NDAEND 110
QY 165 LNSTALGLRTLRLHGVTVSSDVNLNFKDNKGQSSSTANTQIEGEIRGVNLFRASLVAPP 224
Db 111 LYTTSIRFRLLRHGHVNSCDVFNKFKQGNFKSS-----VTSVDRGLLELYQASVLRVH 166
QY 225 GEKVMDEATFSTKYLREALQKIPASSILSLERDVLVYGWHTNLPRLEARNYMDVFGQH 284
Db 167 GEDILDEALSTFTHHLSLAVASLDYP--LSEEVSHALKQSIIRGLPRVEARHYSVY-QD 223

```



```

RESULT 11
DCSI GOSHI STANDARD; PRT; 554 AA.
AC P93665;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE (+)-delta-cadinene synthase (EC 4.2.3.13) (D-cadinene synthase).
GN CDN1.
OS Gossypium hirsutum (Upland cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
ON NCBI_TaxID=3635;
RX [1]
RN PARTIAL SEQUENCE, AND CHARACTERIZATION.
RP MEDLINE=96351891; PubMed=8728715;
RC STRAIN=cv, Westburn M;
RA Davis E.M., Chen Y.-S., Essenberg M., Pierce M.L.;
RT "cDNA sequence of a (+)-delta-cadinene synthase gene induced in
RT Gossypium hirsutum L. by bacterial infection.";
RN [2]
RN PARTIAL SEQUENCE, AND CHARACTERIZATION.
RP MEDLINE=96351891; PubMed=8728715;
RA Davis E.M., Tsuji J., Davis G.D., Pierce M.L., Essenberg M.;
RT "Purification of (+)-delta-cadinene synthase, a sesquiterpene cyclase
RT from bacteria-inoculated cotton foliar tissue.";
RL Phytochemistry 41:1047-1055(1996).
CC (-) FUNCTION: Responsible for the cyclization of trans,trans-farnesyl
CC diphosphate (FPP) to (+)-delta cadinene.
CC CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+) -
CC delta-cadinene + diphosphate.
CC (-) PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
CC first (committed) step.
CC (-) INDUCTION: BY bacterial infection.
CC (-) SIMILARITY: Belongs to the terpene synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U88318; AAC12784.1; -
CC HSSP; Q40577; 5EAU.
CC InterPro; IPR008930; Terp_cyc_toroid.
CC InterPro; IPR001906; Terp_synth-like.
CC InterPro; IPR005630; Terpene_synth_C.
CC InterPro; IPR008949; Terpenoid_synth.
CC Pfam; PF01397; Terpene_synth_1.
CC Pfam; PF03936; Terpene_synth_C_1.
CC
CC ACT_SITE 451 BY SIMILARITY.
CC FT ACT_SITE 527 BY SIMILARITY.
CC FT ACT_SITE 531 BY SIMILARITY.
CC SEQUENCE 554 AA; 64019 MW; 8BC78AD8CA5B816 CRC64;
CC
Query Match 21.3%; Score 692.5; DB 1; Length 554;
Best Local Similarity 31.1%; Pred No. 2,3e-39;
Matches 179; Conservative 113; Mismatches 241; Indels 43; Gaps 15;
QY 51 VASTOSVORRVGNYSHNLWDDDFIQSLISTPYGAPD-----YRERADRLIGEYKIDMFN 104
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
14 LSSNKDEMRPKADFQPSIWGDFL-----NCPDKNDAETAQKRHQQLKEEVKMI-- 63
QY 105 FKSLEDGGNDLQRLLLVDDVERLIGIDRHFYKEIKTALDYNSYWNKIGICGRESVYTD 164
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
64 ----VAMPANSTL-KLAFIDSVQGLGVSHVFTEIEDELE--NIVN-----NNDAEND 110
QY 165 LNSTALGLRTLRLHGYTVSSDVLNVFKDKNGOFSTANIOIEGEIRGVNLFRASLVAPP 224

```

```
Db 111 LYTTSLRPLRLREHGFHVSQVFNKFDQGNFKSS-----VTSQVGRGLLELYQASLYLRVH 166
Qy 225 GEKVMDEATFTSKYLREALQKIPASSITLSLEIRDLVLEYGWHTNLPRLAARNYMDVFGQH 284
Db 167 GEDILDEAISFTSNHLSLAVASL--DHPLSEEVSHALKQSIRRGUPRVEARHLYSVY-QD 223
Qy 285 TKNKNAAEKLELAKLEFNIFHSLQERELKHSVRWKD--SGSPENMTFCRHRHVEYYALAS 343
Db 224 IESHN--KVLLEFAKIDFNWQLLHKKELSEISRWKDLDFQKLPYARDRVVEGYFWLS 281
Qy 344 CIAFPQSGRGLGTTKMSHLITVLDMDYVGVGTVDLELFTATIKRWDPSAMECLPEYM 403
Db 282 GYFEPQYSLGRKMLTKVIAMASIVDDTYDSVATYEELIPYTNALRWMDIKCIDELPEYM 341
Qy 404 KGVYMMVYHTVNEARVAEKAQGRDTLNYARQAEACFDSYMOEAKWIATGYPTEEYL 463
Db 342 KPSYKALLDVVEEMQQLVAEHGRQRYVEYAKNMTIRLAQSYLVEARWTLQNYKPSFEFK 401
Qy 464 ENGVSSAHRPCALQPIILT-DIPFPDHILKEVDPPSKLNDLICILRLRGDTRCYKADR 522
Db 402 ANALTCGYAMLAITSFVGMGDIPTPETPKAANDP-KIIQASTIICRFMDVDVTEHKEKH 460
Qy 523 ARGEBASSICVMKNDPGLTEBDALNHFMTRDAIRELNWELLKPDNSVPITSKKHAPD 582
Db 461 REDDCSAECYMEEY-GVTAQEAQVDFNKHVESAWKQVNOGFLKP-TEMPTEVLNRSLN 518
Qy 583 IGRVWHGGRYRDGYSFANVETKSLVMRTVIEPVPL 618
Db 519 LARVMDVLRGDDGYTYGKAAKGITSLIEPIAL 554

RESULT 12
CASS_RICCO STANDARD; PRT; 601 AA.
ID CASS_RICCO
AC P59287;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Casbene synthase, chloroplast precursor (EC 4.2.3.8).
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalyphaceae;
OC Ricinus.
OX NCBI_TaxID=3988;
RN [1]_TaxID=3988;
RP SEQUENCE FROM N.A.
RX MEDLINE=94359958; PubMed=8078910;
RA Mau C.J., West C.A.;
RT "Cloning of casbene synthase cDNA: evidence for conserved structural
RT features among terpenoid cyclases in plants.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:8497-8501(1994).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=97115639; PubMed=8954576;
RA Hill A.M., Cane D.E., Mau C.J., West C.A.;
RT "High level expression of Ricinus communis casbene synthase in
RT Escherichia coli and characterization of the recombinant enzyme.";
RL Arch. Biochem. Biophys. 336:283-289(1996).
CC -!- FUNCTION: Catalyzes the cyclization of geranylgeranyl diphosphate
CC to casbene, a diterpene phytoalexin with antibacterial and
CC antifungal activity.
CC -!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = casbene +
CC diphosphate.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- INDUCTION: By oligogalacturonide fragments released by fungal
CC infection. Detected after 5 h of incubation with the pectic
CC fragments and reaches a maximum after 10-12 h.
CC -!- MISCELLANEOUS: The Km of this enzyme is 1.9 micromol.
CC -!- SIMILARITY: Belongs to the terpene synthase family.
CC -!- CAUTION: It is uncertain whether Met-1 or Met-8 is the initiator.
-----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on way
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; L32134; -; NOT ANNOTATED CDS.
DR InterPro; IPR008930; Terp_cyc_toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR InterPro; IPR005630; Terpene synth_C.
DR InterPro; IPR008949; Terpenoid synth.
DR Pfam; PF03936; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth_C; 1.
KW Plant defense; Lyase; Chloroplast; Transit peptide.
FT CHAIN 1 56 CHLOROPLAST (POTENTIAL).
FT ACT_SITE 500 500 CASBENE SYNTHASE.
FT ACT_SITE 575 575 BY SIMILARITY.
FT ACT_SITE 579 579 BY SIMILARITY.
SQ SEQUENCE 601 AA; 68965 MW; F7B362D286747957 CRC64;

Query Match 20.5%; Score 667; DB 1; Length 601;
Best Local Similarity 29.8%; Pred. No. 1.3e-37;
Matches 174; Conservative 132; Mismatches 230; Indels 48; Gaps 19;

Qy 47 CLTSTAASDVQRVGNHNSLWDDDTQISLTPTGAPDYERADRLIGEVDKIMFNPK 106
Db 54 CLSS--TTHQEVRLAYPEPTWGNRF-ASLTNFPSEPSYDERVILVKKVKDILISST 110
Qy 107 SLEGGNDLLQRLVDDVERLIGIDRHFKKEIKTALDYV-NS---YWNKSGTCCGRESV 162
Db 111 S-----DSVETVILDLCLRGVSYHFENDIEELLSKIFNSQPLDVBK-----E 155
Qy 163 TDLNSTALGLRLHLGYTVSSDVLNVFKDKNGQFSSTANIQIEGIRGLVNLFRASLVA 222
Db 156 CDLYTAAIVFRVFRQGFQKSSDVSFKDKSOGKFKES----LRGDAKWLSLFEASHLS 211
Qy 223 FPGKVMDEATFTSKYL-EALQKIPASSIILSLEIRDLVLEYGWHTNLPRLAARNYMDV 281
Db 212 VHGEDILEEAFATKDYQLQSSAVELFPN--LKRHITNALEQPFHSGVPRLEARKFDILY 268
Qy 282 GQHTKNNAEKLLELAKLEFNIFHSLQERELKHSVRWKDQG-SPEMTFCRHRHVEYYA 340
Db 269 EADIECRN--ETLLEFAKLDYNRVQLLHQQLCQFSKWKMDLNLASDIPYARDRMAEIFF 326
Qy 341 LASCIAFEPQSGRGLGTTKMSHLITVLDMDYDVGTVDELELFTATIKRWDPSAMECLP 400
Db 327 WAVAMTFEPDYATRMIIAKVLLISLIDTIDATVATMEETHILAEAVARNWMSCLEKLP 386
Qy 401 EYKGYMMVYHTVNEARVAEKAQGRD-TLNYARQAEACFDSYMOEAKWIATGYPPTF 459
Db 387 DYMKVIYKLLNTFTSEFEXEL-TAEKGSYSVKYGREAFQELVGRGYLYEAVWRDEKIPSF 445
Qy 460 BEYLENGKVSSAHRPCALQPIILTIDPFPDHI--LKE--VDPPSKLNDLICILRLGD 514
Db 446 DDLYLNGSMTT-----GLPLVSTASPMVGQETGLNEFQWLETNPKLSVASGAFIKLVND 500
Qy 515 TRCYKADRRARGEASSISICVMKNDPGLTEBDALNHFMTRDAIRELNWELLKPDNSVPI 574
Db 501 LTHSVTEQQRGHVASCIDCYMQQH-GVSKDEAVKILQKMATDCKWKEINEECMR-QSQVSV 558
Qy 575 TSKKHAFDISRVWHGGRYRDGYSFANVETKSLVMRTVIEPVPL 618
Db 559 GHLMRIVNLARLTDVSYKYGDGYTDSQ-QLKQPVKGLFVDPISI 601

RESULT 13
SEAS_TOBAC
ID SEAS_TOBAC STANDARD; PRT; 548 AA.
AC Q40577;
DT 15-JUL-1998 (Rel. 36, Created)
```

DT 15-JUL-1998 (Rel. 36, last sequence update)  
 DE 15-MAR-2004 (Rel. 43, last annotation update)  
 DE Aristolochene synthase (BC 4.2.3.9) (5-epi-aristolochene synthase)  
 (EAS).  
 OS Nicotiana tabacum (Common tobacco).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 CC lamids; Solanales; Solanaceae; Nicotiana.  
 CX NCBI\_TaxID=4097;  
 RN [1]  
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 56-73.  
 RC STRAIN=CV. NK326;  
 RX MEDLINE=93068390; PubMed=1438319;  
 RA Facchini P.J., Chappell J.;  
 RT "Gene family for an elicitor-induced sesquiterpene cyclase in  
 tobacco";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:11088-11092(1992).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS), AND REVISIONS.  
 RX MEDLINE=97442533; PubMed=9295271;  
 RA Starks C.M., Back K., Chappell J., Noel J.P.;  
 RT "Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-  
 aristolochene synthase";  
 RL Science 277:1815-1820(1997).  
 CC -!- FUNCTION: Catalyzes the cyclization of trans,trans-farnesyl  
 diphosphate (FPP) to the bicyclic intermediate 5-epi-  
 aristolochene, initial step in the conversion of FPP to the  
 sesquiterpenoid antifungal phytoalexin capsidiol.  
 CC -!- CATALYTIC ACTIVITY: Trans,trans-farnesyl diphosphate =  
 aristolochene + diphosphate.  
 CC -!- COFACTOR: Binds 3 magnesium ions per subunit.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- INDUCTION: By fungal elicitor.  
 CC -!- SIMILARITY: Belongs to the terpene synthase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L04680; AAA19216.1; --  
 DR PIR; T03714; T03714.  
 DR PDB; 5EAS; 15-OCT-97.  
 DR PDB; 5EAT; 12-NOV-97.  
 DR PDB; 5EAU; 08-APR-98.  
 DR InterPro; IPR008930; Terp\_cyc\_toroid.  
 DR InterPro; IPR001906; Terp\_synth-like.  
 DR InterPro; IPR005630; Terpene synth.C.  
 DR InterPro; IPR008949; Terpenoid synth.  
 DR Pfam; PF01397; Terpene synth; 1.  
 DR Pfam; PF03936; Terpene synth.C; 1.  
 KW Lyase; Magnesium; 3D-structure.  
 FT ACT\_SITE 273  
 FT ACT\_SITE 444 444  
 FT ACT\_SITE 520 520  
 FT ACT\_SITE 525 525  
 FT CONFLICT 42 42  
 FT CONFLICT 44 44  
 FT CONFLICT 55 55  
 FT CONFLICT 62 62  
 FT CONFLICT 73 73  
 FT CONFLICT 89 89  
 FT CONFLICT 388 388  
 FT HELIX 26 29  
 FT HELIX 36 57  
 FT TURN 58 58  
 FT TURN 60 61  
 FT HELIX 64 76  
 FT TURN 77 78  
 Y -> Y IY (IN REF. 1).  
 K -> Q (IN REF. 1).  
 N -> S (IN REF. 1).  
 M -> R (IN REF. 1).  
 T -> I (IN REF. 1).  
 D -> E (IN REF. 1).  
 T -> M (IN REF. 1).

FT HELIX 80 83  
 FT HELIX 84 97  
 FT HELIX 104 116  
 FT TURN 117 118  
 FT HELIX 123 129  
 FT STRAND 130 130  
 FT TURN 132 133  
 FT STRAND 136 136  
 FT HELIX 138 142  
 FT HELIX 144 154  
 FT TURN 155 156  
 FT TURN 159 160  
 FT HELIX 162 164  
 FT TURN 165 166  
 FT HELIX 167 178  
 FT HELIX 179 181  
 FT TURN 184 185  
 FT HELIX 186 195  
 FT TURN 199 200  
 FT HELIX 203 213  
 FT TURN 214 214  
 FT HELIX 215 217  
 FT TURN 219 220  
 FT HELIX 223 254  
 FT TURN 255 255  
 FT HELIX 256 259  
 FT TURN 261 262  
 FT HELIX 267 277  
 FT HELIX 281 283  
 FT HELIX 284 306  
 FT TURN 307 307  
 FT HELIX 310 322  
 FT TURN 323 323  
 FT HELIX 326 330  
 FT TURN 333 354  
 FT TURN 355 357  
 FT HELIX 359 361  
 FT HELIX 362 385  
 FT TURN 386 386  
 FT HELIX 391 398  
 FT TURN 399 399  
 FT HELIX 400 402  
 FT HELIX 404 413  
 FT TURN 414 414  
 FT TURN 416 417  
 FT HELIX 420 427  
 FT TURN 428 428  
 FT HELIX 431 454  
 FT TURN 455 456  
 FT TURN 458 459  
 FT HELIX 461 469  
 FT TURN 470 470  
 FT HELIX 473 494  
 FT TURN 495 495  
 FT HELIX 503 505  
 FT TURN 506 506  
 FT HELIX 507 519  
 FT TURN 520 520  
 FT HELIX 535 542  
 FT TURN 543 543  
 SQ SSSEQUENCE 548 AA; 62973 MW; 9FE1C59CF1A68BF1 CRC64;

Query Match 19.8%; Score 644.5; DB 1; Length 548;  
 Best Local Similarity 28.6%; Pred. No. 3.9e-36;  
 Matches 167; Conservative 122; Mismatches 249; Indels 45; Gaps 14;  
 QY 46 MCLTSVAS-TDSVQRVGVNHSNLDLDDFTQSLISTFYGAPDYRERADRIGVKKIMFN 104  
 Db 1 MASAAVANYEEIEIVRFVADFPSPSLMGDQFLSFDINQV-AEKYAKBIKALKKEQTRNML- 58  
 QY 105 FKSLDDGGNDLLQRLLLVDDVERLGDTRHFKKEIKTALDYVNSYWNKEKGICGGSVWTD 164  
 Db 59 -----ATGMKLADTLNLDITIERLGISYHFEKIDILDOIYN-----QNSCND 103

KW Ligase; Ubl conjugation pathway. 37C2E1BCA0803268 CRC64;  
SQ SEQUENCE 1941 AA; 223682 MW; 37C2E1BCA0803268 CRC64;  
Query Match 4.1%; Score 133; DB 1; Length 1941;  
Best Local Similarity 20.3%; Pred. No. 0.81;  
Matches 141; Conservative 85; Mismatches 244; Indels 226; Gaps 34;  
QY 104 NFKSLDEGG-----NDLLQRL--LVD---DYERLQID-----R 132  
DB 176 NCKGAEDNGRLEDEDDHDKGKISXMLESVLIEFLDFHFDVFNQNIETPTTIQKELIAKL 235  
QY 133 HFKE-----LKTALDYVNSYMKKIGCGRESVVTDLNSTALGLRTL---RLHG 180  
DB 236 YFNEREYEQADMRLRLAYRNQWDEESNKRHLTSLDPLSTLKDVAILVYDEPHY 295  
QY 181 TVSSDLNVFKXGQFS-STANIQIEGEIRGVNLFRASLVAFPG----- 225  
DB 296 SOASAAIRGGPDNKHIDLLTAKIDSEG--RSLLRC-SADIASLMGRIFSQSNGLSCTI 352  
QY 226 ---EKWDEAETSTKYLREALQKIPASSILSL-----EI 257  
DB 353 TQWYEYLHQEACKYSIMWINDCL-NIPNSTFQSIFRNAIGKVLCSKYEPFQSIDMTSV 411  
QY 258 RDVLEYGWHHTNLPRLARNYMDVFGQHTKNKNAEKLLE-----LAKLEFNIFHS 307  
DB 412 RYFSDSYLSDDPYLYADH--SVLGEVGIPLGRHKSLDPGDISAISILNKVIAEDHHE 469  
QY 308 LQERELKHV---SRWKD-----SGS--- 325  
DB 470 YNSRLQYVLFLENRYWKLRKIVQDLIPTLASSAVQKPMFTDQVLEIFPHMTRSGTFM 529  
QY 326 ---PEMTTCRRHHVEYALASCIAPESQHSRGLFTKMSHLITVLDMDYDVGTVDELE 382  
DB 530 DREPQLTSLRESWQLFTCPT-TAYSIFHS-----HFNYLINSVIDVDFSNDEGT 582  
QY 383 LFTATIKRWDPS-----AMECL-----PEYM--KGVMYMYV---HTVNEMAR 419  
DB 583 LVMQVQSNPSKYSISFKQGLYAVETLLSKITPNLLKPGEFIMVTLCKLFGAMK 642  
QY 420 VAEKAQGRDT-----LNARQAWB--ACFDSYQWBAK-----WTATGYLPTFEY 462  
DB 643 IKKKEGHEVLRDQHFIPYLEYTTTSVYSIIQTFDKVQQSKDHIDQRLILGAINLDSF 702  
QY 463 LENGKVS-----SAHRPCALQILTLDIPPDHILKEVDFP--SKLNDLIC 506  
DB 703 LGRNLSYKLYKDFEIKFQISKEQVSNPWNPHLFCFLVQHVPLQVSIQVLSQSKDYL 762  
QY 507 II-LRLRGDTRCYKADRA---RGEASSISCYMKDNPGI---TEEDALNHNFMIT--RD 556  
DB 763 ISDFALRSVVLCSQIDIGFVNRGMSVLHQSAYYKKNPEMSSYSRDIQLNQLAFLEKND 822  
QY 557 AIREL-----NWELLK-PDNSVPITSKHAFDISRV 586  
DB 823 FQVIYNMLDRWELLDFDGSVPSTETVYDDKISSI 858  
RESULT 15  
BXA2\_CLOBO STANDARD; PRT; 1295 AA.  
ID BXA2\_CLOBO  
AC Q45894; P77780;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BoNT/A)  
DE (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-chain; Botulinum neurotoxin A, heavy-chain].  
DE BOT A OR BNA OR ATX.  
GN Clostridium botulinum.  
OS Clostridiales; Clostridia; Clostridiales; Clostridiaceae;  
OC Bacteria; Firmicutes; Clostridia;  
OC Clostridium.  
NCBI\_TaxID=1491;  
RN [1]  
RP SEQUENCE FROM N.A.

165 LNSTALGLRTLRHGYTVSSDLNVFKDNGQFSSSTANIQIEGIRGVNLFRASLVAF 224  
DB 104 LCTSAOLFRLLRQGFNLSPEIFSKQDENGKFKES-----LASDVLGLNLYEASHVR 159  
QY 225 GEKVMDEAETFTSKYLREALQKIPASSILSLEIRDLVEYGHMTNLPRLARNYMDVFG 284  
DB 160 ADDILEDALAFSTTHLESAAPHL--KSPRLQVTHALEQCLHKGVPVRETFISSIYDK 217  
QY 285 TKNKNAEKLLEAKLEBPNIFSHQERELKHVSRWKD--SGSPMTFCRRHHVEYALAS 343  
DB 218 EQSKN--NVLRFKALDNLLQMLHQLKQELAQVSRWKKDLDFVTLPYARDRVCEYFAL 275  
QY 344 CIAEPQHSGRPLGFTKMSHLITVLDMDYDVGTVDELELFTATIKRWDPSAMECLPEYM 403  
DB 276 GYVEPQVSQARVMVKTISMSIVDDTFDAYGVTKVELEYDAIQWDEIDRLPDYM 335  
QY 404 KGVMYMYVHTVNEARVAEKAQGRDTLNARQAWBACFDSYQWBAKWTATGYLPTFEYL 463  
DB 336 KISYKAILDLVKDYKEKELSSAGRSIVCHAIERKMEVVYNNVESTFWIEGYTPPVSYL 395  
QY 464 ENKVSSAHRPCALQILTLDIPPDHILKEVDFP-----SKNLDLCIILRLRGDTRCY 518  
DB 396 SNALATYYIYLAATSYLGM-----KQATEQDFEWSKNPKILEASVILCKVIDDTATY 449  
QY 519 KADRARGEASSISCYMKDNPGIETEDALNHNFMIRDAIRLNLWELLKPNDSVPITSK- 577  
DB 450 EVEKSRGGIATGIECMRDY-GISTYKEMAKFQNAETAWKDINEGLLRP---TPVSTEF 505  
QY 578 --KHAFDISRWVHHGYR--DGYSFANVETKSLVMTVIEPVLP 618  
DB 506 LTPILNLAIRIVEVTYIHNLDGVTPEKVLKPHIINLLVDSIKI 548  
RESULT 14  
UBR1\_KLULA STANDARD; PRT; 1941 AA.  
ID UBR1\_KLULA  
AC O60014;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE N-end-recognizing protein (Ubiquitin-protein ligase E3 component) (N-recogin).  
DE recognin).  
GN UBR1.  
OS Kluyveromyces lactis (Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.  
NCBI\_TaxID=28985;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Waller P.R.H., Varshavsky A.;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Recognition component of the N-end rule pathway. Binds to proteins bearing amino-terminal residues that are destabilizing according to the N-end rule, but does not bind to otherwise identical proteins bearing stabilizing amino-terminal residues.  
CC -!- SIMILARITY: Contains 1 UBR1-type zinc finger.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF061554; AAC15841.1; -.  
DR PIR; T30554; T30554.  
DR InterPro; IPR003126; Znf\_Nrecogin.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF02207; zf-UBR1; 1.  
DR SMART; SM00184; RING; 1.  
DR SMART; SM00396; Znf\_UBR1; 1.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:39:32 ; Search time 45 Seconds

(without alignments)  
4333.117 Million cell updates/sec

Title: US-10-025-145A-65

Perfect score: 3251

Sequence: 1 MALLSIRPLVRSCLSSSHE.....FANVETKSLVMRTVIEPVPL 618

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3251	100.0	618	10	Q948Z0 abies grand
2	2416	74.3	623	10	Q94FW0 abies grand
3	2328.5	71.6	629	10	Q84KL6 pinus taeda
4	2266.5	69.7	627	10	Q84KL4 pinus taeda
5	2264	69.6	630	10	Q9M7D1 abies grand
6	2237	68.8	634	10	Q94KA5 picea abies
7	2205.5	67.8	633	10	Q94KA4 picea abies
8	2163.5	66.5	627	10	Q84SM8 picea abies
9	2151	66.2	628	10	Q84KL3 pinus taeda
10	2137	65.7	630	10	Q9M7D0 abies grand
11	2084.5	64.1	637	10	Q9M7C9 abies grand
12	2041.5	62.8	637	10	Q94FV9 abies grand
13	1985.5	61.1	615	10	Q84KL2 pinus taeda
14	1909	58.7	574	10	Q84KL5 pinus taeda
15	1295.5	39.8	581	10	O64404 abies grand
16	1242.5	38.2	579	10	Q94FW3 abies grand

17	1240	38.1	577	10	Q94XA3
18	1231	37.9	593	10	O64405
19	1187	36.5	782	10	Q9SAU6
20	1187	36.5	817	10	O81086
21	1178.5	36.3	816	10	Q94FW2
22	1131	34.8	873	10	Q947C4
23	1055.5	32.5	853	10	Q94FW1
24	1052.5	32.4	868	10	Q38710
25	836	25.7	603	10	Q8GUE4
26	824	25.3	606	10	O8LSK3
27	821	25.3	606	10	O8LSK1
28	812	25.0	597	10	Q93X23
29	796.5	24.5	595	10	Q9AR86
30	794.5	24.4	595	10	Q7XAS7
31	765.5	23.5	606	10	Q84LB4
32	757	23.3	576	10	O84LB2
33	749.5	23.1	582	10	Q94G53
34	747.5	23.0	607	10	O8WJ39
35	744	22.9	613	10	Q940E7
36	742	22.8	583	10	Q7Y1V1
37	736.5	22.7	598	10	O81192
38	732	22.5	583	10	Q9SPN1
39	728.5	22.4	567	10	Q9SPN0
40	727	22.4	620	10	Q9FUW5
41	724.5	22.3	599	10	Q40322
42	721	22.2	597	10	O8LSJ7
43	717.5	22.1	599	10	Q9SW76
44	710.5	21.9	551	10	Q9LKN1
45	710.5	21.9	603	10	O04806

## ALIGNMENTS

### RESULT 1

Q948Z0	PRELIMINARY;	PRT;	618 AA.
ID	Q948Z0		
AC	Q948Z0;		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)		
DE	(-) - camphene synthase.		
GN	AG6.5.		
OS	Abies grandis (Grand fir).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Coniferopsida; Pinaceae; Pinaceae; Abies.		
OX	NCBI_TaxID=46611;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99373092; PubMed=9268308;		
RA	Bohlmann J., Steele C.L., Croteau R.;		
RT	"Monoterpene synthases from grand fir (Abies grandis). cDNA isolation, characterization, and functional expression of myrcene synthase, (-)-(4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase."		
RT	J. Biol. Chem. 272:21784-21792(1997).		
RL	[2]		
RN	SEQUENCE FROM N.A.		
RP	MEDLINE=99373092; PubMed=10441373;		
RX	Bohlmann J., Phillips M., Ramachandiran V., Katoh S., Croteau R.;		
RT	"cDNA cloning, characterization, and functional expression of four new monoterpene synthase members of the Tpsd gene family from grand fir (Abies grandis)."		
RT	Arch. Biochem. Biophys. 368:232-243(1999).		
RL	EMBL; U87910; AAB70707.1;		
DR	GO; GO:0016829; F-lyase activity; IEA.		
DR	GO; GO:0008152; P-metabolism; IEA.		
DR	InterPro; IPR005630; Terpene synth. C.		
DR	InterPro; IPR008949; Terpenoid synth.		
DR	InterPro; IPR008930; Terp. cyc. toroid.		
DR	InterPro; IPR001906; Terp. synth.-like.		
DR	Pfam; PF01397; Terpene synth; 1.		
DR	Pfam; PF03936; Terpene synth C; 1.		
DR	SEQUENCE 618 AA; 70749 MW; B07B5185CE5C4CE1 CRC64;		
DR	SEQUENCE 618 AA; 70749 MW; B07B5185CE5C4CE1 CRC64;		

Query Match 100.0%; Score 3251; DB 10; Length 618;  
 Best Local Similarity 100.0%; Pred. No. 9e-232;  
 Matches 618; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALLSITPLVSRCLSSSHEIKALRTPITLIGICRPGKSAHNSINMCLTSVASTDSVQRR 60  
 DB 1 MALLSITPLVSRCLSSSHEIKALRTPITLIGICRPGKSAHNSINMCLTSVASTDSVQRR 60

QY 61 VGNHNSLWDDDFIQSLISTPYGADYRERADRLIGEVKDIMFNFKSLDGGNDLQRL 120  
 DB 61 VGNHNSLWDDDFIQSLISTPYGADYRERADRLIGEVKDIMFNFKSLDGGNDLQRL 120

QY 121 LVDDVERLIGIDRHFKKEIKTALDYVNSYWNKGGICGRESVVTDLNSALGLRLT 180  
 DB 121 LVDDVERLIGIDRHFKKEIKTALDYVNSYWNKGGICGRESVVTDLNSALGLRLT 180

QY 181 TVSSDVLNVFKDKNGQFSSTANI QIEGIRGVNLFRASLVAFPGCKYMDAEFTSTKYL 240  
 DB 181 TVSSDVLNVFKDKNGQFSSTANI QIEGIRGVNLFRASLVAFPGCKYMDAEFTSTKYL 240

QY 241 REALQKIPASSILSLEIRDLVEYGMHTNLPRLARNYMDVFGQHTKNKAAEKLLEAKL 300  
 DB 241 REALQKIPASSILSLEIRDLVEYGMHTNLPRLARNYMDVFGQHTKNKAAEKLLEAKL 300

QY 301 EFNIFHSLQERELKHVRWKKDGSPEMTFCRHRHVEYYALASCIAPQSGFRLGFTK 360  
 DB 301 EFNIFHSLQERELKHVRWKKDGSPEMTFCRHRHVEYYALASCIAPQSGFRLGFTK 360

QY 361 MSHLITVLDDMYDVFGTVDELELTATIKWDPSCAMECLPEYMGVVMVYHTVNEARV 420  
 DB 361 MSHLITVLDDMYDVFGTVDELELTATIKWDPSCAMECLPEYMGVVMVYHTVNEARV 420

QY 421 AEKQAGRDITLNYARQAEACFDSYMQEAKWIATGYLTFEYLENGKVSAAHRCALQPI 480  
 DB 421 AEKQAGRDITLNYARQAEACFDSYMQEAKWIATGYLTFEYLENGKVSAAHRCALQPI 480

QY 481 LTLDIPPDHILKEVDPPSKNDLICIILRLRGDTRCYKADRAGEASSISYCKMDNPG 540  
 DB 481 LTLDIPPDHILKEVDPPSKNDLICIILRLRGDTRCYKADRAGEASSISYCKMDNPG 540

QY 541 LTBEDALNHNFMIRDAIRELNWELLKPDNSVPTTSKKHAFDISRVWHGYYRVDGYSFA 600  
 DB 541 LTBEDALNHNFMIRDAIRELNWELLKPDNSVPTTSKKHAFDISRVWHGYYRVDGYSFA 600

QY 601 NVETKSLVMTVIEPVPL 618  
 DB 601 NVETKSLVMTVIEPVPL 618

RESULT 2  
 Q94FWO PRELIMINARY; PRT; 623 AA.

ID Q94FWO;  
 AC Q94FWO;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Pinene synthase (Fragment).  
 OS Abies grandis (Grand fir).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.  
 OX NCBI\_TaxID=46611;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21297238; PubMed=11404343;  
 RA Trapp S.C., Croteau R.B.;  
 RT "Genomic organization of plant terpene synthases and molecular  
 evolutionary implications";  
 RL Genetics 158:811-832(2001).  
 DR EMBL; AF326517; AAK83564.1; --  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR005630; Terpene\_synth\_C.

DR InterPro; IPR008949; Terpenoid\_synth.  
 DR InterPro; IPR008930; Terp\_cyc\_coroid.  
 DR InterPro; IPR001906; Terp\_synth-like.  
 DR Pfam; PF01397; Terpene\_synth; 1.  
 DR Pfam; PF03936; Terpene\_synth\_C; 1.  
 FT NON TER  
 SQ SEQUENCE 623 AA; 70948 MW; 823B82B38113467B CRC64;

Query Match 74.3%; Score 2416; DB 10; Length 623;  
 Best Local Similarity 74.8%; Pred. No. 5.2e-170;  
 Matches 467; Conservative 59; Mismatches 82; Indels 16; Gaps 6;

QY 8 PLVSRCLSSSHEIKALRTPITLIGICRPGKSAHNSINMCLTSVASTDSVQRRVG 62  
 DB 3 PLAKSCLHKLSSSHEIKALRTPITLIGICRPGKSAHNSINMCLTSVASTDSVQRRVG 62

QY 63 NYHNSLWDDDFIQSLISTPYGADYRERADRLIGEVKDIMFNFKSLDGGNDLQRL 117  
 DB 63 NYHNSLWDDDFIQSLISTPYGADYRERADRLIGEVKDIMFNFKSLDGGNDLQRL 117

QY 118 LVDDVERLIGIDRHFKKEIKTALDYVNSYWNKGGICGRESVVTDLNSALGLRLT 177  
 DB 121 LVDDVERLIGIDRHFKKEIKTALDYVNSYWNKGGICGRESVVTDLNSALGLRLT 177

QY 178 HGYTVSSDVLNVFKDKNGQFSSTANI QIEGIRGVNLFRASLVAFPGCKYMDAEFTST 237  
 DB 181 HGYTVSSDVLNVFKDKNGQFSSTANI QIEGIRGVNLFRASLVAFPGCKYMDAEFTST 237

QY 238 KYREALQKIPASSILSLEIRDLVEYGMHTNLPRLARNYMDVFGQHTKNKAAEKL 294  
 DB 241 KYREALQKIPASSILSLEIRDLVEYGMHTNLPRLARNYMDVFGQHTKNKAAEKL 294

QY 295 LEIHLAKLQERELKHVRWKKDGSPEMTFCRHRHVEYYALASCIAPQSGF 354  
 DB 300 LEIHLAKLQERELKHVRWKKDGSPEMTFCRHRHVEYYALASCIAPQSGF 354

QY 355 RLQFTKMSHLITVLDDMYDVFGTVDELELTATIKWDPSCAMECLPEYMGVVMVYHTV 414  
 DB 360 RLQFTKMSHLITVLDDMYDVFGTVDELELTATIKWDPSCAMECLPEYMGVVMVYHTV 414

QY 415 NEMARVAEQAQGRDITLNYARQAEACFDSYMQEAKWIATGYLTFEYLENGKVSAAHRC 474  
 DB 420 NEMARVAEQAQGRDITLNYARQAEACFDSYMQEAKWIATGYLTFEYLENGKVSAAHRC 474

QY 475 CALQPIITLDIPPDHILKEVDPPSKNDLICIILRLRGDTRCYKADRAGEASSISY 534  
 DB 480 CALQPIITLDIPPDHILKEVDPPSKNDLICIILRLRGDTRCYKADRAGEASSISY 534

QY 535 MKDNPGLTBEDALNHNFMIRDAIRELNWELLKPDNSVPTTSKKHAFDISRVWHGYYR 594  
 DB 540 MKDNPGLTBEDALNHNFMIRDAIRELNWELLKPDNSVPTTSKKHAFDISRVWHGYYR 594

QY 595 DGYSFANVETKSLVMTVIEPVPL 618  
 DB 600 DGYSFANVETKSLVMTVIEPVPL 618

RESULT 3  
 Q84KL6 PRELIMINARY; PRT; 629 AA.

ID Q84KL6;  
 AC Q84KL6;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE (-)-alpha-pinene synthase.  
 OS Pinus taeda (loblolly pine).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.  
 OX NCBI\_TaxID=3352;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22510022; PubMed=12623076;  
 RA Phillips M.A., Wildung M.R., Williams D.C., Hyatt D.C., Croteau R.;



```
ID Q9W7D1 PRELIMINARY; PRT; 630 AA.
AC Q9W7D1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Beta-phellandrene synthase.
GN AGC8.
OS Abies grandis (Grand fir).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Pinaceae; Abies.
OX NCBI_TaxID=46611;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stem;
RX MEDLINE=93373092; PubMed=10441373;
RA Bohlmann J., Phillips M., Ramachandiran V., Katoh S., Croteau R.;
RT "cDNA cloning, characterization, and functional expression of four new
RT monoterpene synthase members of the Tpsd gene family from grand fir
RT (Abies grandis).";
RL Arch. Biochem. Biophys. 368:232-243 (1999).
DR EMBL; AF139205; AAF61453.1; -.
DR HSP; Q40577; 5EAB.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp cyc toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth C; 1.
DR PFam; PF03936; Terpene synth C; 1.
SQ SEQUENCE 630 AA; 72784 MW; BBE4374B262FF2D1 CRC64;

Query Match 69.68; Score 2264; DB 10; Length 630;
Best Local Similarity 68.38; Pred. No. 9.2e-159;
Matches 435; Conservative 78; Mismatches 98; Indels 26; Gaps 9;

QY 1 MALLSITPLVSRSL-----SSSHEIKALRRITPTLIGICRPGKSAHSINMLCITSVAST 54
DB 1 MALVSSAP---KSLHKSLIRSTHHEKLPRLRTPTLGMCRGKSFPSVMSLTTAVSD 57
QY 55 DSVQRRVGNVHSLNWDHDFQSLISTPYGADPYRERADRLIGEVKDIMFNPKSLDGG-- 112
DB 58 DGLQRRIGDGHSLNWDHDFQSL-STPYGEPYSRERAEKLGIVKE-MFNSMPEDEGSM 115
QY 113 ---NDLLQRLLLVDVVERLGIDRHFKEIKTALDYVNSYWNKEGICGGRSVVTDLNSTA 169
DB 116 SPNDLIERLWVDSVERLGIDRHFKEIKSALDYVSYWNKEGICGGRSDVPFDVNSTA 175
QY 170 LGLRTLRHGYTVSSDVLNFKDKNGQFSSTANIQIEGIRGVNLFRASLVAPPGEKVM 229
DB 176 SGFRTLHGYTVSSSEVLKVFQDQGFAPSPSTK-ERDITVLNLYRASFIAPPGKVM 234
QY 230 DEARTSTKYLREALOKIPASSILSLRIDVLEVGWHTNLPRLRNMYDMVFQGH--- 285
DB 235 EEAEIFSSRYLKGAQVQIPVSS-LSQEIDVTLEYGHTNPRLETRNYLDVFGHTSPWL 293
QY 286 KKNKA----AEKLELAKLEFNIFHSLQERELKHVSRRWKDGSPEMTFCRRHVEYYAL 341
DB 294 KKQRTQYLDSEKLELAKLEFNIFHSLQERELKHVSRRWKDGSPEMTFCRRHVEYYAL 353
QY 342 ASCIAFEPQHSRGLGFTKMSHLITVLDMDYVFGTVDELELFTATIKRWDPDSAMECLPE 401
DB 354 SSCIAFEPKHSARFLGPAKTHLITVLDMDYVFGTVDELELFTATIKRWDPDSAMECLPE 413
QY 402 YMGKVYVMVYHTVNMARVAEKAQGRDITLNYARQAEACFDSYMOEAKWTATGYLPTFEE 461
DB 414 YMEIYALYALTDMAEAEKTCGRDITLNYARQAEACFDSYMOEAKWTATGYLPTFEE 473
QY 462 YLENGKVSSAHRPCALQIITLDIPDPHILKEVDFPSKLDNLICILRLRGDTRCYKAD 521
DB 474 YLENKGVSSGHRAAALTPLLTLDVPLDDVLKGDIPSPRENDLASSFLRLRGDTRCYKAD 533
QY 522 RARGEAEASSISYKONPGITEEDALNHINFMIDALRELNWELLKPDNSVPTTSKKHAF 581
```

```
DB 534 RDRGEAEASSISYKONPGITEEDALNHINFMIDALRELNWELLKPDNSVPTTSKKHAF 593
QY 582 DISRWHHGYRDRDGYSFANVETKSLVMRTVIEPVL 618
DB 594 EITRAFHLQYKRDGFSVATQETKSLVRRTVLEPVL 630

RESULT 6
Q94KA5 PRELIMINARY; PRT; 634 AA.
AC Q94KA5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Beta-phellandrene synthase-like protein.
OS Picea abies (Norway spruce) (Picea excelsa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Pinaceae; Picea.
OX NCBI_TaxID=3329;
RN [1]
RP SEQUENCE FROM N.A.
RA Morency M.J., Nicole M.C., Seguin A.;
RT "Terpene synthase from Norway spruce, cDNA isolation and
RT characterization of beta-phellandrene synthase-like gene.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF169918; AAK39127.2; -.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp cyc toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth C; 1.
DR PFam; PF03936; Terpene synth C; 1.
SQ SEQUENCE 634 AA; 72768 MW; BE140A49C46B4404 CRC64;

Query Match 68.88; Score 2237; DB 10; Length 634;
Best Local Similarity 68.08; Pred. No. 9.2e-157;
Matches 434; Conservative 75; Mismatches 105; Indels 24; Gaps 10;

QY 1 MALLSITPLVSRSL-----SSSHEIKALRRITPTLIGICRPGKSA-HSINMLCIT-SVAS 53
DB 1 MSPSVIPLAYKLCPLPSLSSSREVKPLHTIPTNLGMCRRGKSMAPASTSMILTAASVD 60
QY 54 TDSVQRRVGNVHSLNWDHDFQSLISTPYGADPYRERADRLIGEVKDIMFNPKSLDGG- 112
DB 61 DSVQRRVGNVHSLNWDHDFQSL-STPYGEPYSRERAEKLGIVKE-MFNSIKDDEL 118
QY 113 ---NDLLQRLLLVDVVERLGIDRHFKEIKTALDYVNSYWNKEGICGGRSVVTDLNST 168
DB 119 ITPNDLIERLWVDSVERLGIDRHFKEIKSALDYVSYWNKEGICGGRSDVVADLNST 178
QY 169 ALGLRTLHGYTVSSDVLNFKDKNGQFSSTANIQIEGIRGVNLFRASLVAPPGEKV 228
DB 179 ALGLRTLHGYTVSSSEVLKVFQDQGFAPSPSTK-KTEGIRGALNLYRASLIAPPGKV 237
QY 229 MDEARTSTKYLREALOKIPASSILSLRIDVLEVGWHTNLPRLRNMYDMVFQGH--- 285
DB 238 MDDEAFSSRYLKGAQVQIPVSS-LSQEIDVTLEYGHTNPRLETRNYLDVFGHTSPW 296
QY 286 --KKNKA----AEKLELAKLEFNIFHSLQERELKHVSRRWKDGSPEMTFCRRHVEYYA 340
DB 297 LKKNKTYMDGEKLELAKLEFNIFHSLQERELKHVSRRWKDGSPEMTFCRRHVEYYT 356
QY 341 LASCIAFEPQHSRGLGFTKMSHLITVLDMDYVFGTVDELELFTATIKRWDPDSAMECLP 400
DB 357 LGSCIAFEPKHSARFLGPAKTHLITVLDMDYVFGTVDELELFTATIKRWDPDSAMECLP 416
QY 401 EYMGKVYVMVYHTVNMARVAEKAQGRDITLNYARQAEACFDSYMOEAKWTATGYLPTFE 460
DB 417 DYMGKVYVMVYALTEMAQEAQKTCGRDITLNYARQAEACFDSYMOEAKWTATGYLPTFE 476
```



Db 238 EAEIFSATYLKEALQIPVSS-LSQEIQVYLQVRMHSNLPRLRLEARTYIDILQENTKNOML 236  
 QY 291 ---AEKLELAKLEFNIHSLQERELKHVSRRWKDGSPEMTFCRHRHVEYYALASCIAP 347  
 Db 297 DVNTKKVLELAKLEFNIHSLQERELKHVSRRWKESGFPDLNFIHRHVEFYTLVSGIDM 356  
 QY 348 EPHSGFRGLGFTKMSHLITVLDDMYDVFQTVDELELFTATIKRWDPSSAMECLPEYMKGV 407  
 Db 357 EPHGCTFRLSFVKMCHLITVLDDMYDTFGTIDELRLFTAARKWDPSTTECLPEYMKGV 416  
 QY 408 MMVYHTVNMARVAEKAQGRDTLNYARQAEACFDSYMOBAKIATGYLPTPEYLENGK 467  
 Db 417 TWLYETVNMARQAEKQSGRDTLSYVRQALEAVIGAYHKEABEVISSGYLPTPEYFENGK 476  
 QY 468 VSSAHRPCALQILTLIDIPPPDHILKEVDPPSKLNDLICIILRLRGDTRCYKADRRAGEE 527  
 Db 477 VSSGHRIATLQPTFMJIDIPPPHVLQEIIDPPSKFNDIPFACSLRLRGDTRCYQADRRAGEE 536  
 QY 528 ASSISCYMKDNPGLTEEDALNHINFMIRDAIRELNWELLKPDNSVPITSKKHAFDLSRVW 587  
 Db 537 ASSISCYMKDNPQSTQEDALNHINNMLEETIKKLNWELLKPDNNVPISSKKHAFDINRGL 596  
 QY 588 HHGYRDRGYSFANVETKSLVMRTVIEPVPL 618  
 Db 597 HHFYNYRDRGYTVASNETKNLVIKTVLEPVPM 627

## RESULT 9

Q84KL3 PRELIMINARY; PRT; 628 AA.  
 ID Q84KL3; AC Q84KL3;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE (+)-alpha-pinene synthase.  
 OS Pinus taeda (Loblolly pine).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.  
 OX NCBI\_TaxID=3352;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22510022; PubMed=12623076;  
 RA Phillips M.A., Wildung M.R., Williams D.C., Hyatt D.C., Croteau R.;  
 RT "cDNA isolation, functional expression, and characterization of (+)-  
 alpha-pinene synthase and (-)-alpha-pinene synthase from loblolly pine  
 (Pinus taeda): Stereocontrol in pinene biosynthesis.";  
 RL Arch. Biochem. Biophys. 411:267-276 (2003).  
 DR EMBL; AF543530; AAO61228.1; -;  
 DR GO; GO:0016829; P:lyase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR005630; Terpene synth. C.  
 DR InterPro; IPR008949; Terpenoid synth.  
 DR InterPro; IPR008930; Terp. cyc. toroid.  
 DR InterPro; IPR001906; Terp. synth.-like.  
 DR Pfam; PF01397; Terpene synth; 1.  
 DR Pfam; PF03936; Terpene synth C; 1.  
 SQ SEQUENCE 628 AA; 71495 MW; 4208222612256837 CRC64;

Query Match 66.2%; Score 2151; DB 10; Length 628;  
 Best Local Similarity 65.8%; Pred. No. 2.1e-150;  
 Matches 416; Conservative 92; Mismatches 106; Indels 18; Gaps 8;

QY 1 MALLSTPLVSRSL-----SSSHEIKALRRITPTLIGICRPGKSVAHSNM-CLTSVAST 54  
 Db 1 MALVSAPLNSKLCRLRTLFGFSHELKAIHSTVPLNGLMCRGGKSIAPSSMSTTSVSNE 60  
 QY 55 DSQVRRVGNVHNSLNWDDDFQTSILSTPYGAPDYRRERADRLIGVVKDIMEFNFKSLEDCG-- 112  
 Db 61 DGVPRRIAGHSNLWDDDSIASL-STSYEAPSYRKRADKLIGVKNI-FDLMSVEDGVFT 118  
 QY 113 ---NDLLQLLLVDDVVERLIGIDRHFKEIKTALDYNSVYNEKIGCGRESVYTDLNSTA 169  
 Db 119 SPLSDLHRLWMDVVERLIGIDRHFKEIKTALDYNSVYNEKIGCGRESVYTDLNSTA 178

QY 170 LGLRTLRLHGYTVSSDVLNVFQDKNGQFQSTANIQIEGEIRGVNLPRASLVAFPPGKVM 229  
 Db 179 LGLRTLRLHGYTVSSHLVDHFKNEKQFTCSA-IQTEGEIRVNLPRASLVAFPPGKIM 237  
 QY 230 DEATEFSTKYLRALOKIPASSILSLERDVLEVGHNTLPRLEARNYMDVFOGHTKNKN 289  
 Db 238 EAAEIESFYLKDALQKIPPSG-LSQEIYLLFVFGWHTNLPRLMETRMVYIDVFGEDITFET 296  
 QY 290 ---AAEKILELAKLEFNIHSLQERELKHVSRRWKDGSPEMTFCRHRHVEYYALASCIA 346  
 Db 297 PYLIREKLELAKLEFNIHSLQERELKHVSRRWKDYGPPETFSRHRHVEYYTLAACA 356  
 QY 347 FEPOHSGFRGLGFTKMSHLITVLDDMYDVFQTVDELELFTATIKRWDPSSAMECLPEYMKGV 406  
 Db 357 NDKHSAFRLGFKISHMITLDDIYDTGTMEELKLLTAAFKRWDPSSIECLPDYMKGV 416  
 QY 407 YMVYHTVNMARVAEKAQGRDTLNYARQAEACFDSYMOBAKIATGYLPTPEYLENG 466  
 Db 417 YXAVYDNINEMAREAKIQGWDTVSYARKSWEAFIGAYIQEAKWISSGYLPTPEYLENG 476  
 QY 467 KUSSAHRPCALQILTLIDIPPPDHILKEVDPPSKLNDLICIILRLRGDTRCYKADRRAGE 526  
 Db 477 KVSFGSRITTLPEMLTLGFPPLPRILQEIIDFPKFNDLICAILRLKGDTCYKADRRAGE 536  
 QY 527 EASSISCYMKDNPGLTEEDALNHINFMIRDAIRELNWELLKPDNSVPITSKKHAFDLSRV 586  
 Db 537 EASAVSCYMKDHPGITEEDAVNQVNAVNDNLKELNELLRPDSGVPISYKKAADFICRV 596  
 QY 587 HHGYRDRGYSFANVETKSLVMRTVIEPVPL 618  
 Db 597 HHGYRDRGYSFASIEIKNLVTRTVETVPL 628

## RESULT 10

Q9MW7D0 PRELIMINARY; PRT; 630 AA.  
 ID Q9MW7D0; AC Q9MW7D0;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Terpinolene synthase.  
 GN AGC9.  
 OS Abies grandis (Grand fir).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.  
 OX NCBI\_TaxID=46611;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Stem;  
 RX MEDLINE=99373092; PubMed=10441373;  
 RA Bohlmann J., Phillips M., Ramchandiran V., Katoh S., Croteau R.;  
 RT "cDNA cloning, characterization, and functional expression of four new  
 monoterpene synthase members of the Tpsd gene family from grand fir  
 (Abies grandis).";  
 RL Arch. Biochem. Biophys. 368:232-243 (1999).  
 DR EMBL; AF139206; AAF61454.1; -;  
 DR HSSP; Q40577; 5EAU.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR005630; Terpene synth. C.  
 DR InterPro; IPR008949; Terpenoid synth.  
 DR InterPro; IPR008930; Terp. cyc. toroid.  
 DR InterPro; IPR001906; Terp. synth.-like.  
 DR Pfam; PF01397; Terpene synth; 1.  
 DR Pfam; PF03936; Terpene synth C; 1.  
 SQ SEQUENCE 630 AA; 72508 MW; 452437B87F203D8A CRC64;

Query Match 65.7%; Score 2137; DB 10; Length 630;  
 Best Local Similarity 67.4%; Pred. No. 2.3e-149;  
 Matches 427; Conservative 73; Mismatches 114; Indels 20; Gaps 9;

QY 1 MALLSTPLVSRSL-----SSSHEIKALRRITPTLIGICRPGKSVAHSNMCLTSVASTD 55

```

Db 1 MALVSIILPLSKSVLHKSIIWVSTYEHKAIISITIPNLGRGKSVTHSLRSLSTAVSDD 60
QY 56 -SVQRRVGVNHSNLDWDDFIQSLISTPYGAPDYERADRILGEVNDIMFNFKSLEDGG-- 112
Db 61 HGVQRRIVEFHNSLWDDDFIQL-STPYGASYSRERADRLIVEVKG-FTSISAEDGELI 118
QY 113 ---NDLLQRLLLVDDVERLIGDRHFKKIKTALDYVNSYWNKGIQCGRESVWTDLNSTA 169
Db 119 TPLNDLIQRLMVDNVERLIGDRHFKKIKTALDYVNSYWNKGIQCGRESVWTDLNSTA 178
QY 170 LGLRLRLHGYTVSSDVLNFKD--KNGFSTANIQTGIRGVNLNFRASLAFPGCK 227
Db 179 LGRILRLHGYTVSSDVLNFKD--KNGFSTANIQTGIRGVNLNFRASLAFPGCK 237
QY 228 VMDAEFTSTKYREALQKIPASSILSLERDVLVLEYGWHTNLPRLAARNYMDVFGHTKN 287
Db 238 VMEAEFTSTKYREALQKIPASSILSLERDVLVLEYGWHTNLPRLAARNYMDVFGHTKN 296
QY 288 KNA---AEKLLLEAKLEFNIFHSLOERELKHVSRWKKDGSPEMTFCRRHVEYVALASC 344
Db 297 ETLVNMKLELLEAKLEFNIFHSLOERELKHVSRWKKDGSPEMTFCRRHVEYVALASC 356
QY 345 IAFEPHSGFRGLFTKMSHLITVLDMDYDVGTVDELELTATIKRWDPSAWECLPEYMK 404
Db 357 IETDRKSGFRGLFTKMSHLITVLDMDYDVGTVDELELTATIKRWDPSAWECLPEYMK 416
QY 405 GYVMVYHTVNMARVAEKAQGRDITLNYARQAWACFDSYMOEAKIATGYLPTFEYILE 464
Db 417 GLYVMVYHTVNMARVAEKAQGRDITLNYARQAWACFDSYMOEAKIATGYLPTFEYILE 476
QY 465 NGKVSAAHRPCALQIPILTDIPFPDHILKEVDFFSKNDLICIILRLGRDTRCYKADRAR 524
Db 477 TSKVSFGYRIEALQIPILTDIPFPDHILKEVDFFSKNDLICIILRLGRDTRCYKADRAR 536
QY 525 GEEASSICYMKDNPGLTEEDALNHNFMIRDAIRELNWELLKPDNSVPIITSKKHAFDIS 584
Db 537 GEEASSICYMKDNPGLTEEDALNHNFMIRDAIRELNWELLKPDNSVPIITSKKHAFDIS 596
QY 585 RVWHGGRYRDGYGFANVETKSLVMRTVIEFVPL 618
Db 597 KGSLLHGYKYRDFGFSVANKETKNWVRTVLESVPL 630

```

## RESULT 11

```

ID Q9M7C9 PRELIMINARY; PRT; 637 AA.
AC Q9M7C9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE (-)-limonene/(-)-alpha-pinene synthase.
GN AGC11.
OS Abies grandis (Grand fir).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_taxID=46611;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Stem;
RX MEDLINE=99373092; PubMed=10441373;
RA Bohlmann J., Phillips M., Ramachandiran V., Katoh S., Croteau R.;
RT "cDNA cloning, characterization, and functional expression of four new
RT monoterpene synthase members of the Tpsd gene family from grand fir
RT (Abies grandis).";
RL Arch. Biochem. Biophys. 368:232-243 (1999).
DR EMBL; AF139207; AAF61455.1; -.
DR HSPF; Q40577; SEAS.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cyc_Teroid.

```

```

DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene synth. 1.
DR Pfam; PF03936; Terpene synth C; 1.
SQ SEQUENCE 637 AA; 73273 MW; B3574986FEC96CFB CRC64;

Query Match 64.1%; Score 2084.5; DB 10; Length 637;
Best Local Similarity 64.9%; Pred. No. 1.8e-145;
Matches 417; Conservative 80; Mismatches 115; Indels 31; Gaps 13;

QY 1 MALLSIIVLQVPCGGLKSLISSNVQKALCISTAVPLRMRKQKALV--INKKLITVS 52
Db 1 MALLSIIVLQVPCGGLKSLISSNVQKALCISTAVPLRMRKQKALV--INKKLITVS 58
QY 53 STDS-----VQRRVGVNHSNLDWDDFIQSLISTPYGAPDYERADRILGEVNDIMFN- 105
Db 59 HRDNGGGVQLQRIADHHPNLWEDDFIQSL--SSPYGSSYSERAVTVVEVKE-MFNSIP 116
QY 106 --KSLDGGNDLQRLLLVDDVERLIGDRHFKKIKTALDYVNSYWNK-GIGCGRESV 162
Db 117 NNRELFSGNDLLTFLMWMVDSIERLIGDRHFKKIKTALDYVNSYWNK-GIGCGRESV 176
QY 163 TDLNSTALGLRLHLGYTVSSDVLNFKD--KNGFSTANIQTGIRGVNLNFRASL 221
Db 177 PDINSTALALTLRLHGVNSDVLVLEYGKQGHFACPA--ILTEGQITRSVNLVRS 235
QY 222 APPEKVMDEATSTKYREALQKIPASSILSLERDVLVLEYGWHTNLPRLAARNYMDV 281
Db 236 APPEKVMDEATSTKYREALQKIPASSILSLERDVLVLEYGWHTNLPRLAARNYMDV 294
QY 282 GQHTKNKN-----AAEKLLEAKLEFNIFHSLOERELKHVSRWKKDGSPEMTFCRRH 335
Db 295 GQSVESNEMPYNTQKLLAKLEFNIFHSLOERELKHVSRWKKDGSPEMTFCRRH 354
QY 336 VEYVALASCIAPFQHSGRGLFTKMSHLITVLDMDYDVGTVDELELTATIKRWDPSA 395
Db 355 VEYITMASCISWEPKHSARLFGVKTCHLLVLDMDYDVGTVDELELTATIKRWDPSA 414
QY 396 MECLPEYMKGYVMVYHTVNMARVAEKAQGRDITLNYARQAWACFDSYMOEAKIATGY 455
Db 415 TKCLPEYMKGYVMVYHTVNMARVAEKAQGRDITLNYARQAWACFDSYMOEAKIATGY 474
QY 456 LPTFEVLENGKVSAAHRPCALQIPILTDIPFPDHILKEVDFFSKNDLICIILRLGR 515
Db 475 LPTFEVLENGKVSAAHRPCALQIPILTDIPFPDHILKEVDFFSKNDLICIILRLGR 534
QY 516 RCYKADRRARGEASSISCYMKDNPGLTEEDALNHNFMIRDAIRELNWELLKPDNSVPI 575
Db 535 RCYKADRRARGEASSISCYMKDNPGLTEEDALNHNFMIRDAIRELNWELLKPDNSVPI 594
QY 576 SKKHAFDISRVWHGGRYRDGYGFANVETKSLVMRTVIEFVPL 618
Db 595 SKKHAFDISRVWHGGRYRDGYGFANVETKSLVMRTVIEFVPL 637

```

## RESULT 12

```

ID Q94FV9 PRELIMINARY; PRT; 637 AA.
AC Q94FV9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE (-)-4S-limonene synthase.
OS Abies grandis (Grand fir).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_taxID=46611;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21297238; PubMed=11404343;
RA Trapp S.C., Croteau R.B.;
RT "Genomic organization of plant terpene synthases and molecular
RT evolutionary implications.";
RL Genetics 158:811-832 (2001).

```





Search completed: July 7, 2004, 17:43:43  
Job time : 48 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:41:07 ; Search time 52 Seconds  
(without alignments)  
3699.488 Million cell updates/sec

Title: US-10-025-145A-65

Perfect score: 3251

Sequence: 1 MALLSITPLVSRSCSSSHE.....FANVETKSLVMRTVIEVPVL 618

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 31123816 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3251	100.0	618	14	US-10-025-145A-65
2	2453	75.5	628	9	US-09-887-586A-20
3	2453	75.5	628	9	US-09-903-012-20
4	2453	75.5	628	10	US-09-900-797-20
5	2453	75.5	628	12	US-09-893-820-20
6	2453	75.5	628	13	US-10-041-007-22
7	2453	75.5	628	14	US-10-025-145A-4
8	2264	69.6	630	14	US-10-025-145A-67
9	2200.5	67.7	627	9	US-09-887-586A-30
10	2200.5	67.7	627	9	US-09-903-012-30
11	2200.5	67.7	627	10	US-09-900-797-30
12	2200.5	67.7	627	12	US-09-893-820-30
13	2200.5	67.7	627	13	US-10-041-007-26
14	2200.5	67.7	627	14	US-10-025-145A-2
15	2197.5	67.6	627	14	US-10-025-145A-32

16	2137	65.7	630	14	US-10-025-145A-78
17	2084.5	64.1	637	13	US-10-041-007-28
18	2084.5	64.1	637	14	US-10-025-145A-69
19	2031.5	62.5	637	9	US-09-887-586A-58
20	2031.5	62.5	637	9	US-09-903-012-58
21	2031.5	62.5	637	10	US-09-900-797-58
22	2031.5	62.5	637	12	US-09-893-820-58
23	2031.5	62.5	637	13	US-10-041-007-24
24	2031.5	62.5	637	14	US-10-025-145A-6
25	1295.5	39.8	581	9	US-09-887-586A-48
26	1295.5	39.8	581	9	US-09-903-012-48
27	1295.5	39.8	581	10	US-09-900-797-48
28	1295.5	39.8	581	12	US-09-893-820-48
29	1295.5	39.8	581	13	US-10-041-007-18
30	1250.5	38.5	577	14	US-10-025-145A-18
31	1240.5	38.2	862	9	US-09-887-586A-44
32	1240.5	38.2	862	9	US-09-903-012-44
33	1240.5	38.2	862	10	US-09-900-797-44
34	1240.5	38.2	862	12	US-09-893-820-44
35	1240.5	38.2	862	12	US-10-041-018-386
36	1240.5	38.2	862	13	US-10-041-007-41
37	1231	37.9	593	9	US-09-887-586A-50
38	1231	37.9	593	9	US-09-903-012-50
39	1231	37.9	593	10	US-09-900-797-50
40	1231	37.9	593	12	US-09-893-820-50
41	1231	37.9	593	13	US-10-041-007-20
42	1187	36.5	782	9	US-09-887-586A-46
43	1187	36.5	782	9	US-09-903-012-46
44	1187	36.5	782	10	US-09-900-797-46
45	1187	36.5	782	12	US-09-893-820-46

#### ALIGNMENTS

#### RESULT 1

US-10-025-145A-65  
; Sequence 65, Application US/10025145A  
; Publication No. US20030175861A1  
; GENERAL INFORMATION:  
; APPLICANT: Croteau, Rodney B.  
; APPLICANT: Bohlmann, Joerg  
; APPLICANT: Steele, Christopher L.  
; APPLICANT: Phillips, Michael A.  
; TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)  
; FILE REFERENCE: WSUR118414  
; CURRENT APPLICATION NUMBER: US/10/025,145A  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: US 09/360,545  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: PCT/US98/14528  
; PRIOR FILING DATE: 1998-07-10  
; PRIOR APPLICATION NUMBER: US 60/052,249  
; PRIOR FILING DATE: 1997-07-11  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 65  
; LENGTH: 618  
; TYPE: PRT  
; ORGANISM: Abies Grandis  
US-10-025-145A-65

Query Match 100.0%; Score 3251; DB 14; Length 618;  
Best Local Similarity 100.0%; Pred. No. 1.4e-302;  
Matches 618; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MALLSITPLVSRSCSSSHEIKALRRPTPTIGICRPGKSVASHSNMCLTSTVASTDSVQR	60
Db	1	MALLSITPLVSRSCSSSHEIKALRRPTPTIGICRPGKSVASHSNMCLTSTVASTDSVQR	60
Qy	61	VGNVHNSLWDDFTQSLISTPYGAPYRERADRLIGEVKDIMFNFKSLDGGNDLLQRL	120
Db	61	VGNVHNSLWDDFTQSLISTPYGAPYRERADRLIGEVKDIMFNFKSLDGGNDLLQRL	120

```

QY 121 LVDDVERLGDHREKKEIKTALDYVNSYNWNEKIGCGRESVWTDLNSTALGLRTRLHG 180
Db 121 LVDDVERLGDHREKKEIKTALDYVNSYNWNEKIGCGRESVWTDLNSTALGLRTRLHG 180
QY 181 TVSSDVNLVFKKNGQFSSSTANIQIEGEIRGVNLFRASLVAPPGKVMDEAETFTSKYL 240
Db 181 TVSSDVNLVFKKNGQFSSSTANIQIEGEIRGVNLFRASLVAPPGKVMDEAETFTSKYL 240
QY 241 REALQKIPASSILSLSEIRDVLEYGWHTNLPRLERARNYMDVFGQHTKNKNAAEKLELAKL 300
Db 241 REALQKIPASSILSLSEIRDVLEYGWHTNLPRLERARNYMDVFGQHTKNKNAAEKLELAKL 300
QY 301 EFNIFHSLQRELUKHSRWKMGSGSPMTFCRRHVEYVYALASCIAPFPGHSGFRLGFTK 360
Db 301 EFNIFHSLQRELUKHSRWKMGSGSPMTFCRRHVEYVYALASCIAPFPGHSGFRLGFTK 360
QY 361 MSHLITVLDMDYDVFQTVDELELFTATIKRWDPSSAMECLPEYMKGVYMTVYHTVNMARV 420
Db 361 MSHLITVLDMDYDVFQTVDELELFTATIKRWDPSSAMECLPEYMKGVYMTVYHTVNMARV 420
QY 421 AEXAQGRDTLNYARQAWACFDSYMOEAKWIAATGYLPTPEEYLENGKVSSAHRPCALQPI 480
Db 421 AEXAQGRDTLNYARQAWACFDSYMOEAKWIAATGYLPTPEEYLENGKVSSAHRPCALQPI 480
QY 481 LTLDDIPPPHILKEVDPPSKNDLICIILRLRGDTRCYKADRGARGEASISCYMKDNP 540
Db 481 LTLDDIPPPHILKEVDPPSKNDLICIILRLRGDTRCYKADRGARGEASISCYMKDNP 540
QY 541 LTEDALNHINFMIRDAIRELNWELLKPDNSVPIITSKKHAFDISRVWHGYRYRDGYSFA 600
Db 541 LTEDALNHINFMIRDAIRELNWELLKPDNSVPIITSKKHAFDISRVWHGYRYRDGYSFA 600
QY 601 NVETKSLVMRTVIEPVPL 618
Db 601 NVETKSLVMRTVIEPVPL 618

```

RESULT 2

```

US-09-887-586A-20
; Sequence 20, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT FILING DATE: 2001-06-22
; PRIOR FILING DATE: 1999-09-17
; PRIOR FILING DATE: 1999-04-22
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Abies grandis
US-09-887-586A-20

```

Query Match 75.5%; Score 2453; DB 9; Length 628;  
 Best Local Similarity 75.3%; Pred. No. 5.9e-226;  
 Matches 475; Conservative 59; Mismatches 81; Indels 16; Gaps 6;

```

QY 1 MALLSTPLVSRGCL-----SSSHEIKALRRTIPTIGICRPGKSVASHNMLCLTSVASTD 55
Db 1 MALVSTAPLASKCLHKLSSLSSTHELKALSTETIPALCMRGKASITPSISMSSTTVTTDD 60

```

```

QY 56 SVORRVGNVHSLWDDDFIQSLISTPYGAPDYRERADRLIGEVKIDIMFNPKSLEDGG--- 112
Db 61 GVRRRMGDPHSLNWDVVVIQSL-PTAYEKSYLEAEKULIGEVEN-MFNSMSLEDGELMS 118
QY 113 --NDLQRLLLVDDVERLGDHREKKEIKTALDYVNSYNWNEKIGCGRESVWTDLNSTAL 170
Db 119 PLNDLIQRLWI VDSLERLGIHREHFKOEIKSALDYVYVWGENGIGCGRESVWTDLNSTAL 178
QY 171 GLRTILRHGYTVSSDVNLVFKKNGQFSSSTANIQIEGEIRGVNLFRASLVAPPGKVM 230
Db 179 GLRTILRHGYTVSSDVNLVFKKNGQFSSSTANIQIEGEIRGVNLFRASLVAPPGKIMD 238
QY 231 EATFTSTKYLRALQKIPASSILSLSEIRDVLEYGWHTNLPRLERARNYMDVFGQHTKNKNA 290
Db 239 EAEIFSTKYLKEALQKIPVSS--LSREIGDVLEYGWHTYLPRLERARNYI QVFGQDTENTKS 297
QY 291 ---AEKLELAKLEFNIFHSLQRELUKHSRWKMGSGSPMTFCRRHVEYVYALASCIAP 347
Db 298 YVSKKLELAKLEFNIFQSLQKRELESVRWVKESGFPEMTFCRRHVEYVYTLASCIAP 357
QY 348 EPOHSGFRLGFTKMSHLITVLDMDYDVFQTVDELELFTATIKRWDPSSAMECLPEYMKGVY 407
Db 358 EPOHSGFRLGFTKCHLITVLDMDYDVFQTVDELELFTATMKRWDPSSIDCLPEYMKGVY 417
QY 408 MMVYHTVNMARVAEXAQGRDTLNYARQAWACFDSYMOEAKWIAATGYLPTPEEYLENGK 467
Db 418 IAYDVTVNMAREAEAEAQGRDTLYAREAEWYDSYMQEARWIAATGYLPSFDEYENGK 477
QY 468 VSSAHRPCALQPIITLDDIPPPHILKEVDPPSKNDLICIILRLRGDTRCYKADRGAR 527
Db 478 VSCGHRISALQPIITLDDIPPPHILKEVDPPSKNDLACAILRLRGDTRCYKADRGAR 537
QY 528 ASSISCYMKDNPGLTREDALNHINFMIRDAIRELNWELLKPDNSVPIITSKKHAFDISRV 587
Db 538 ASSISCYMKDNPVSEEDALDHINAMISDVIKLANWELLKPDINVPISAKKHAFDIARAF 597
QY 588 HHGYRYRDGYSFANVETKSLVMRTVIEPVPL 618
Db 598 HHGYRYRDGYSVANVETKSLVTRTLESVPL 628

```

RESULT 3

```

US-09-903-012-20
; Sequence 20, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT FILING DATE: 2001-07-11
; PRIOR FILING DATE: 1999-09-17
; PRIOR FILING DATE: 1999-09-17
; PRIOR FILING DATE: 1998-09-18
; PRIOR FILING DATE: 1998-09-18
; PRIOR FILING DATE: 1999-04-22
; PRIOR FILING DATE: 1999-04-22
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Abies grandis
US-09-903-012-20

```

Query Match 75.5%; Score 2453; DB 9; Length 628;  
 Best Local Similarity 75.3%; Pred. No. 5.9e-226;  
 Matches 475; Conservative 59; Mismatches 81; Indels 16; Gaps 6;

```

Qy      1 MALLSITPLVSRCL-----SSSHIKAIRPTIPTGICRPGKSVAHISNNCLTSVASTD 55
Db      1 MALYSTAPLASKSLCHKLSLISSTHELKALSRTPALGMSRRGKSITPSPISMSSTVTVD 60

Qy     56 SVQRGVNGYHNLMDDDDTIOQLSIPTPYGAPDYRERADRLIGVKDIMFNFKXLEDDG--- 112
Db     61 GVRRMGDGFHNLMDDDVIQSL-PTYAEKSYLERAEKLIGEVEN-MFNSMSLEDGLMS 118

Qy    113 --NDLLQRLLVDDVERLIGIDRHFFKEIKTALDYVNSYWNEKGICGGRSVVTDLNSTAL 170
Db   119 PLNDLIQRLWTVDSLERLGIHRHFDEIKSALDYVYSYWGENGICGGRSVVTDLNSTAL 178

Qy    171 GLRTLRLHGTVSSDVLMNVFKDKNGOFISTANIQTGEITRGVLNLFRASLVAPPGEKWD 230
Db   179 GLRTLRLHGYPVSSDVFKAFGQNGQFCSENIQTDEELRGVLNLFRASLIAPFGEKIND 238

Qy    231 BAETFTSKYLREALKOKIPASSILSLIEIRDVLBYGWHTNLPRIEARNYMDVFGQHTKNNA 290
Db   239 BAEFTSTKYLEAKLOKIPVSS--LSREIGDVLEYGWHTYLPRIEARNYIOVFQODTENTKS 297

Qy    291 ---AEKULELAKLEFENIFHSIQERELKHVSRWKDGSGPEMTFCRRHRYVEYVALASCIAF 347
Db   298 YVKSUKLELAKLEFENIFSLOKRELESILRVWKESGPPEMTFCRRHRYVEYTLASCIAF 357

Qy    348 EPOHSGFRPLGTFKMGHLLITVLDMDYDFGTVDLEBELFTATIKRPDPSAMECLPEYMKGY 407
Db   358 EPOHSGFRPLGFAKTCHLLITVLDMDYTDTGTVDLEBELFTATMKRWPDPSSIDCLPEYMKGY 417

Qy    408 MMVYHTVNEMARVAEKAGRDLTNLVARQAWEACFDYSMQEAWIATGYLPTFEELYENCK 467
Db   418 IAVYDTVNEMAREAEAAGRDTLTVAREAWAAYIDSUQOEAWIATGYLPSFDEYYENCK 477

Qy    468 VSSAHRPCALOFLITLDIPFPDHILKEVDFPSKLANDLICIIURLRGDTRCYKADRARGE 527
Db   478 VSCGHRISALQFILTMDFPPDHILKEVDFPSKLANDLCAILURLRGDTRCYKADRARGE 537

Qy    528 ASSISCYMKDNPGLTERRDALNHINPMIDAIBELNWELLKPNSVPITTSKKHAFDISRV 587
Db   538 ASSISCYMKDNPGVSEEUDALDHINAMISDVIKGLNWELLKPDINPFI SAKKHAFDIARAF 597

Qy    588 HHGYRYRDGYGFANVETKSLVMRTVIEPVPL 618
Db   598 HYGKYRDGYGFVANVETKSLVTRTLESVPL 628

```

## RESULT 4

```

US-09-900-797-20
; Sequence 20, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Abies grandis
US-09-900-797-20

```

	Query Match	75.5%; Score 2453; DB 10; Length 628;
	Best Local Similarity	75.3%; Pred. No. 5.9e-226;
	Matches	475; Conservative 59; Mismatches 81; Indels 16; Gaps 6
Qy	1	MALLSIPLVSRSCD-----SSSHEIKALRRTIPTLIGICRPKGSVAHSINMCLTISVASTD 55
Db	1	MAIVYSTAPLASKSLHKLSLISSTHELKALSTIIFALGMSRRGKSITPISMSSTTVTTDD 60
Qy	56	SVQRNVGNYSNLDWDFPIQSILSTPYGAPYRERADRLIGEVDIMPNFKSLEDPGG--- 112
Db	61	GVRRMGDFFHNLDDDDVIQSL-PTAYEEKSYLERAELKIEVKKN-MFNSMSLEDGETLMS 118
Qy	113	--NDLLORLLAVDSVERIGIDRHFEKEIKTALDVVNSYMNKEGTGCGRSVVTDLNSTATL 170
Db	119	PLNDLIQRLMTLVDSLERIGIRHRFDEIKSALDVDVYSWGNGTGCGRESVVTDLNSTATL 178
Qy	171	GLRTLRLHGTYVSSDLNVNFKDNGQFSSTANIQTIEGIRGVNLFRASLVAFPGEKVMD 230
Db	179	GLRTLRLHGYPVSSDVFKAFGQNGQFSCSENIQTDEEIRGVNLFRASLIAFPGEKIMD 238
Qy	231	EAEFTSKYLBEALKQKIPASSIIISLETFDVLVEYGWHHTNLPLEARNYMDVPGOHTKNNA 290
Db	239	EAEIFTSKYLEALKQKIPVSS-LREBIGDVLVEYGWHTYLPLEARNYIQVGQDTENTKS 297
Qy	291	---AEKLLEAKLBFNFIHPSLORELKHVSPPWKDSDSGSPEMTFCFHRHVEYYVALASCIAF 347
Db	298	VYKSKULLEAKLBFNFIQSLOKRELESIVRWKESGPPEMTFCFHRHVEYYTLASCIAF 357
Qy	348	EPQSHGPRLGTRKMSHLITVLDDMYDVPGTVDLELELTATTIKRMDPDSAMECLPEYMKGVY 407
Db	358	EPQSHGSRFLGPAKTCHAITVLDDMYDTPGTVDLELELTATTKRWDPDSSIDCLPEYMKGVY 417
Qy	408	MMVYHTVNEMARVAEKAGQRTDLNVAROAWEACPDPSYMOEAKWTATGYLPFEEYLENGK 467
Db	418	IANTVTVNEMAREABEAQGRTLLYAREAWAAYIDSYMOEARWATATGYLPSEFFEYENGK 477
Qy	468	VSSAHRPCALQPIILTIDIPPDPHILKEYDFFPSKLNLDLICILRLRGDTRCYKADRAGEE 527
Db	478	VSCGHRIQSALQPILTMDIIPFDHILKEYDFFPSKLNLDLACAILRLRGDTRCYKADRAGEE 537
Qy	528	ASSISCYMKONPGITEDALNHINFMIEDATELNWELLKPDNSVPIITSKKHAFDIRSW 587
Db	538	ASSISCYMKONPGVSEEDALDHINAMISDVIKGUNWELLKPDINVPISAKKHAFDIARAF 597
Qy	588	HGKYRYRDGYSFANVETKSLVMRTVIPBPVL 618
Db	598	HGKYRYRDGYSVANVETKSIATRTLLAFSVPL 628

## RESULT 5

```

US/09-893-820-20
; Sequence 20, Application US/09893820
; Publication NO. US20040053386A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20040053386A1, Joseph P.
; APPLICANT: Stars, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20

```

```

; LENGTH: 628
; TYPE: PRT
; ORGANISM: Abies grandis
US-09-893-820-20

Query Match      75.5%; Score 2453; DB 12; Length 628;
Best Local Similarity 75.3%; Pred. No. 5.9e-226;
Matches 475; Conservative 59; Mismatches 81; Indels 16; Gaps 6;

QY 1 MALLSITPLVRSCL-----SSSHEIKALRRITPTLIGICRPGKSVAHNSNMCLTSVASTD 55
Db 1 MALVSTAPLASKSCLHKLSSSTHELKALSRITPALGMSRRGKSITPISMSSTTVVTD 60

QY 56 SVQRRVGNVHSLWDDDFQSLSTPYGAPDYRERADRLIGEVKDIEMFNKSLDGG--- 112
Db 61 GVRNRRMGDFHSLWDDVQSL-PTAYEBSKYLERAELKIGEVKN-MFNMSLDELMS 118

QY 113 --NDLQRLLLVDDVERLIGDRHFKEIKTALDYVNSYWNKEGIGCGRESVVTDLNSTAL 170
Db 119 PLNDLQRLWIVDSLERLGHRRHFKEIKSALDYVYSYWGENGIGCGRESVVTDLNSTAL 178

QY 171 GLRTLRLHGYTVSSDVLNVFKQNGQSFSTANIQIEGIRGVNLFRASLVAFPGEKMD 230
Db 179 GLRTLRLHGYTVSSDVLNVFKQNGQSFSTANIQIEGIRGVNLFRASLVAFPGEKMD 238

QY 231 EAEFTSTKYLREALQKIPASSILSLRDLVLEYGWHTNLPRLRARNYVDFGQHTKNKA 290
Db 239 EAEFTSTKYLREALQKIPASSILSLRDLVLEYGWHTNLPRLRARNYVDFGQHTKNKA 297

QY 291 ---AEKLLLELAKLEFNIHSLQERELKHVSRWKKDSGSPMTFCRRHVEYVALASCI 347
Db 298 YVKSLLLELAKLEFNIHSLQERELKHVSRWKKDSGSPMTFCRRHVEYVALASCI 357

QY 348 EPOHSGFRGLGFTQMSHLITVLDMDYVFGTVDELELFTATIKWDPSAMECLPEYMGVY 407
Db 358 EPOHSGFRGLGFTQMSHLITVLDMDYVFGTVDELELFTATIKWDPSAMECLPEYMGVY 417

QY 408 MMVYHTVNMARVAEKAQGRDTLNYARQAEACFDSYMOEAKWIATGYLPTPEYLENGK 467
Db 418 IAVDYTVNWARAEAEQAQGRDTLTYAREAEWYIDSYMOEAKWIATGYLPTPEYLENGK 477

QY 468 VSSAHRPCALQPLTLDIPFDHILKEVDPPSKLNDLICIILRLGDTRCYKADRAGEE 527
Db 478 VSCGHRISALQPLTLDIPFDHILKEVDPPSKLNDLACAILRLGDTRCYKADRAGEE 537

QY 528 ASSISCYMKDNPGLTEDALNHINFMIRDAIRLAWELLKPDNSVPITSKKHAFDISR 587
Db 538 ASSISCYMKDNPGVSEEDALDHINAMISDVIKGLNELLKPDINVPISAKKHAFDIARAF 597

QY 588 HGYKYRDRGYSFANVETKSLVMTVIEPVPL 618
Db 598 HGYKYRDRGYSVANVETKSLVTTLESVPL 628

```

```

RESULT 6
US-10-041-007-22
; Sequence 22, Application US/10041007
; Publication No. US20020164736A1
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seiichi P.T.
; APPLICANT: Schepmann, Hala G.
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase
; FILE REFERENCE: P02081U51
; CURRENT APPLICATION NUMBER: US/10/041,007
; PRIOR FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/259,881
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Abies grandis

```

```

US-10-041-007-22

Query Match      75.5%; Score 2453; DB 13; Length 628;
Best Local Similarity 75.3%; Pred. No. 5.9e-226;
Matches 475; Conservative 59; Mismatches 81; Indels 16; Gaps 6;

QY 1 MALLSITPLVRSCL-----SSSHEIKALRRITPTLIGICRPGKSVAHNSNMCLTSVASTD 55
Db 1 MALVSTAPLASKSCLHKLSSSTHELKALSRITPALGMSRRGKSITPISMSSTTVVTD 60

QY 56 SVQRRVGNVHSLWDDDFQSLSTPYGAPDYRERADRLIGEVKDIEMFNKSLDGG--- 112
Db 61 GVRNRRMGDFHSLWDDVQSL-PTAYEBSKYLERAELKIGEVKN-MFNMSLDELMS 118

QY 113 --NDLQRLLLVDDVERLIGDRHFKEIKTALDYVNSYWNKEGIGCGRESVVTDLNSTAL 170
Db 119 PLNDLQRLWIVDSLERLGHRRHFKEIKSALDYVYSYWGENGIGCGRESVVTDLNSTAL 178

QY 171 GLRTLRLHGYTVSSDVLNVFKQNGQSFSTANIQIEGIRGVNLFRASLVAFPGEKMD 230
Db 179 GLRTLRLHGYTVSSDVLNVFKQNGQSFSTANIQIEGIRGVNLFRASLVAFPGEKMD 238

QY 231 EAEFTSTKYLREALQKIPASSILSLRDLVLEYGWHTNLPRLRARNYVDFGQHTKNKA 290
Db 239 EAEFTSTKYLREALQKIPASSILSLRDLVLEYGWHTNLPRLRARNYVDFGQHTKNKA 297

QY 291 ---AEKLLLELAKLEFNIHSLQERELKHVSRWKKDSGSPMTFCRRHVEYVALASCI 347
Db 298 YVKSLLLELAKLEFNIHSLQERELKHVSRWKKDSGSPMTFCRRHVEYVALASCI 357

QY 348 EPOHSGFRGLGFTQMSHLITVLDMDYVFGTVDELELFTATIKWDPSAMECLPEYMGVY 407
Db 358 EPOHSGFRGLGFTQMSHLITVLDMDYVFGTVDELELFTATIKWDPSAMECLPEYMGVY 417

QY 408 MMVYHTVNMARVAEKAQGRDTLNYARQAEACFDSYMOEAKWIATGYLPTPEYLENGK 467
Db 418 IAVDYTVNWARAEAEQAQGRDTLTYAREAEWYIDSYMOEAKWIATGYLPTPEYLENGK 477

QY 468 VSSAHRPCALQPLTLDIPFDHILKEVDPPSKLNDLICIILRLGDTRCYKADRAGEE 527
Db 478 VSCGHRISALQPLTLDIPFDHILKEVDPPSKLNDLACAILRLGDTRCYKADRAGEE 537

QY 528 ASSISCYMKDNPGLTEDALNHINFMIRDAIRLAWELLKPDNSVPITSKKHAFDISR 587
Db 538 ASSISCYMKDNPGVSEEDALDHINAMISDVIKGLNELLKPDINVPISAKKHAFDIARAF 597

QY 588 HGYKYRDRGYSFANVETKSLVMTVIEPVPL 618
Db 598 HGYKYRDRGYSVANVETKSLVTTLESVPL 628

```

```

RESULT 7
US-10-025-145A-4
; Sequence 4, Application US/10025145A
; Publication No. US20030175861A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Bohlmann, Joerg
; APPLICANT: Steele, Christopher L.
; APPLICANT: Phillips, Michael A.
; TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)
; FILE REFERENCE: WSUR118414
; CURRENT APPLICATION NUMBER: US/10/025,145A
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 09/360,545
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: PCT/US98/14528
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: US 60/052,249
; PRIOR FILING DATE: 1997-07-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4

```

```
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Abies Grandis
US-10-025-145A-4

Query Match      75.5%; Score 2453; DB 14; Length 628;
Best Local Similarity 75.3%; Pred. No. 5.9e-226;
Matches 475; Conservative 59; Mismatches 81; Indels 16; Gaps 6;

Qy 1 MALLSIPTLVRSCL-----SSSHEIKALRRTIPTLIGICRPGKSVASHINCLTSVASTD 55
Db 1 MALVSTAPLASKSCLHKSLSSTHELKALSTIIFALGMSRRGKSIPTISMSSTVTVTDD 60

Qy 56 SVQRVNGVHNSLWDDDFIQSLISTPYGAPDYERADRLIGEVKDIMFNFKSLEDGG--- 112
Db 61 GVRRMGDHNSLWDDDFIQSL-PTAYEEKSYLERAELIGEVKN-MFNSMSLEDGELMS 118

Qy 113 --NDLLQRLLLVDDVERLIGIDRHFKKEIKTALDYVNSYWNKEGICGGRSVVTDLNSTAL 170
Db 119 PLNDLIQRLWTVDSLERLIGIHRHFKDEIKSALDYVYSYWGNGICGGRSVVTDLNSTAL 178

Qy 171 GLRLRLHGYTVSSDVLNVFKKNGQFSSTANIQTIEGIRGVNLFRASLVAFPGKVM 230
Db 179 GLRLRLHGYTVSSDVLNVFKAFKNGQFSSTANIQTIEGIRGVNLFRASLVAFPGKIM 238

Qy 231 EAFTFTKYLREALQKIPASSILSLRDLVLEYGWHTNLPRLRARNYMDVFGQHTKKN 290
Db 239 EAFTFTKYLREALQKIPVSS-LSREIGDVLGYWHTVLPRLRARNYIQVFGQDTEHTKS 297

Qy 291 ---AEKLELAKLEFNIHSLQERELKHVSWMKDSGSPENTFCRHRHVEYYALASCI 347
Db 298 YVSKKLELAKLEFNIHSLQERELKHVSWMKDSGSPENTFCRHRHVEYYALASCI 357

Qy 348 EPQSGFRLGFTKMSHLITVDDMDYDVTGTVDELELFTATIKRWDPPSAMECLPEYMGV 407
Db 358 EPQSGFRLGFTKMSHLITVDDMDYDVTGTVDELELFTATIKRWDPPSAMDCLPEYMGV 417

Qy 408 MMVYHTVNMARVAEAKQRTLNVARQAWACFDSYMQEAKWIATGYLPTFEYLENGK 467
Db 418 IAVYDVTNMARVAEAKQRTLNVARQAWACFDSYMQEAKWIATGYLPTFEYLENGK 477

Qy 468 VSSARPCALPILTDIPFPDHILKEVDFPSKNDLICIILRLRGDTRCYKADRGEE 527
Db 478 VSCGHRISALPILTDIPFPDHILKEVDFPSKNDLICALILRLRGDTRCYKADRGEE 537

Qy 528 ASSISCYMKDNPGLTEEDALNHINFMIRDAIRELNWELLKPDNSVPIITSKKHAFDISRW 587
Db 538 ASSISCYMKDNPGLTEEDALNHINFMIRDAIRELNWELLKPDINVPISAKKHAFDIARAF 597

Qy 588 HHGYRDRGYSFANVETKSLVMRTVIEPVPL 618
Db 598 HHGYRDRGYSFANVETKSLVTRTILLESVPL 628

RESULT 8
US-10-025-145A-67
; Sequence 67, Application US/10025145A
; Publication No. US20030175861A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Bohlmann, Joerg B.
; APPLICANT: Steele, Christopher L.
; APPLICANT: Phillips, Michael A.
; TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)
; FILE REFERENCE: WSUR118414
; CURRENT APPLICATION NUMBER: US/10/025,145A
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 09/360,545
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: PCT/US98/14528
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: US 60/052,249
; PRIOR FILING DATE: 1997-07-11
```

```
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Abies Grandis
US-10-025-145A-67

Query Match      69.6%; Score 2264; DB 14; Length 630;
Best Local Similarity 68.3%; Pred. No. 8.2e-208;
Matches 435; Conservative 78; Mismatches 98; Indels 26; Gaps 9;

Qy 1 MALLSIPTLVRSCL-----SSSHEIKALRRTIPTLIGICRPGKSVASHINCLTSVAST 54
Db 1 MALVSSAP---KSLHKSLSIRSTHELKPLRRTIPTLIGICRGRGKSTFTPSVSMSTLTVASD 57

Qy 55 DSVQRVNGVHNSLWDDDFIQSLISTPYGAPDYERADRLIGEVKDIMFNFKSLEDGG--- 112
Db 58 DGLQRRIGDHSNLDWDDFIQSL-STPYGEPSYERAEKELIGEVK-MFNSMPSDEGESM 115

Qy 113 ---NDLLQRLLLVDDVERLIGIDRHFKKEIKTALDYVNSYWNKEGICGGRSVVTDLNSTA 169
Db 116 SPLNDLIERLMMWDSVERLIGIDRHFKKEIKSALDYVYSYWNKEGICGGRSDVFPDVSNTA 175

Qy 170 LGLRLRLHGYTVSSDVLNVFKKNGQFSSTANIQTIEGIRGVNLFRASLVAFPGKVM 229
Db 176 SGFTLRLHGYTVSSDVLNVFKQDQNGQFAPSPSTK-ERDITVNLNLYRASFIAPFGKVM 234

Qy 230 DEATFTSTKYLREALQKIPASSILSLRDLVLEYGWHTNLPRLRARNYMDVFGQHT--- 285
Db 235 EAEIFSSRYLKEAVQKIPVSS-LSQEDYLTLEYGWHTNMPRLRTRNLDVFGHTSPWL 293

Qy 286 KKNKA----AEKLELAKLEFNIHSLQERELKHVSWMKDSGSPENTFCRHRHVEYYAL 341
Db 294 KKKRTQVLDSEKLELAKLEFNIHSLQERELKHVSWMKDSGSPENTFCRHRHVEYYAL 353

Qy 342 ASCIAPFQSGFRLGFTKMSHLITVDDMDYDVTGTVDELELFTATIKRWDPPSAMECLPE 401
Db 354 SSCIATEPKHSAFRLGFTKMSHLITVDDMDYDVTGTVDELELFTATIKRWDPPSAMECLPE 413

Qy 402 YMKGYVMVYHTVNMARVAEAKQRTLNVARQAWACFDSYMQEAKWIATGYLPTFE 461
Db 414 YMKGYVMVYHTVNMARVAEAKQRTLNVARQAWACFDSYMQEAKWIATGYLPTFE 473

Qy 462 YLENGKYSAAHRCALPILTDIPFPDHILKEVDFPSKNDLICIILRLRGDTRCYKAD 521
Db 474 YLENAKYSAGHRAALPILTDIPFPDHILKEVDFPSKNDLICIILRLRGDTRCYKAD 533

Qy 522 RARGEASSISCYMKDNPGLTEEDALNHINFMIRDAIRELNWELLKPDNSVPIITSKKHAF 581
Db 534 RARGEASSISCYMKDNPGLTEEDALNHINFMIRDAIRELNWELLKPDNSVPIITSKKHAF 593

Qy 582 DISRVWTHGYRDRGYSFANVETKSLVMRTVIEPVPL 618
Db 594 EITRAFQLKYRDRGYSFANVETKSLVTRTILLESVPL 630

RESULT 9
US-09-887-586A-30
; Sequence 30, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
```

; PRIOR FILING DATE: 1999-04-22  
 ; PRIOR APPLICATION NUMBER: 60/150,262  
 ; PRIOR FILING DATE: 1999-08-23  
 ; NUMBER OF SEQ ID NOS: 58  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 30  
 ; LENGTH: 627  
 ; TYPE: PRT  
 ; ORGANISM: Abies grandis  
 ; US-09-887-586A-30

Query Match 67.7%; Score 2200.5; DB 9; Length 627;  
 Best Local Similarity 68.1%; Pred. No. 1e-201;  
 Matches 427; Conservative 77; Mismatches 106; Indels 17; Gaps 7;

QY	1	MALLSITPLVSRSL-----SSSHEIKALRRRTIPTIGICRPGKSVAHNSNMCLTSVASTD	55
DB	1	MAUVSISPLASKCLRKSLISSIHEHKPPYRTIPNIGMRRRGKSVTPSPMSISLATAADD	60
QY	56	SVORRNVGNSHNLWDDDFIQSLISTPYGADPYRERADRLIGEVKDIMFNFKSLEDGG--	112
DB	61	GVORRIGDYSNINWDDDFIQSL-STPYGEPYQERAEERLIVEVKI-FNSMYLDDGRLMS	118
QY	113	--NDLLQRLVDDVERLIGDRHFKKEIKETALDYVNSYNEKIGGCGRESVVTDLNSTAL	170
DB	119	SPNDLMQRLWIVDSVERLGARHFKNEITSALDYVFRYWEENGIGCGRDSIVTDLNSTAL	178
QY	171	GLRTLRHGVTVSSDVLNVFKDNGQFSSTANTQIEGEIRGVNLFRASIVAPPGKVM	230
DB	179	GFETLRHGVTVSPVVKAFQDQNGQFVCSPG-QTEGEIRSVNLFRASIVAPPGKVM	237
QY	231	EATFTSKYLREALQKIPASSILSLRDLVLEYGWHNLPRLRARNYMDVFGQHTK--N	287
DB	238	EABIFSTRYLKEALQKIPVSA-LSQBIKFVMEYGMHNLPRLEARNYMDVFGQHTK--N	296
QY	288	KNAEKLLELAKLEFNI FHSLORELKHVSRRWKDGSPEMTFCRRHRHVEYYALASCI	347
DB	297	KNAGKKLLELAKLEFNI FNSLOQKELQYLLRMWKESDLPKLTFAHRHVEFFYTLASCI	356
QY	348	EPHSGFRGLGFTKMSHLITVDDMDYVGTVDLELFTATIKWDPSAMECLPEYMGVY	407
DB	357	DPKHSFRGLGFTKMSHLITVDDMDYVGTVDLELFTATIKWDPSAMECLPEYMGVY	416
QY	408	MMVYHTVNEARVAEKAQGRDTLNYARQAEACFDYSYMOEAKWATGYLPTFPEYLENGK	467
DB	477	VSSAYRATLQPTLTNLAWLPDYILKGIDFPSPFNDLASSFLRLGRDTRCYKADRGE	536
QY	528	ASSISCYMDNPGCLTEEDALNHINFMIRDAIRLNWELLKPNDSVPTTSKKHAFDTSRV	587
DB	537	ASCISCYMDNPGSTEDALNHINAMVNDIILKELNELLRSNDNIPMLAKKHAFDITRAL	596
QY	588	HGGRYRDGYSFANVETKSLVMRTVIE	614
DB	597	HHLYIYRDGFSVANKETKLVMTLE	623

RESULT 10  
 US-09-903-012-30  
 ; Sequence 30, Application US/09903012  
 ; Patent No. US20020094557A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chappell, Joseph  
 ; APPLICANT: Starks, Courtney M.  
 ; APPLICANT: Manna, Kathleen R.  
 ; TITLE OF INVENTION: SYNTHASES  
 ; FILE REFERENCE: 07678-025001  
 ; CURRENT APPLICATION NUMBER: US/09/903,012  
 ; CURRENT FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: 09/398,395  
 ; PRIOR FILING DATE: 1999-09-17  
 ; PRIOR APPLICATION NUMBER: 60/100,993  
 ; PRIOR FILING DATE: 1998-09-18  
 ; PRIOR APPLICATION NUMBER: 60/130,628  
 ; PRIOR FILING DATE: 1999-04-22  
 ; PRIOR APPLICATION NUMBER: 60/150,262  
 ; PRIOR FILING DATE: 1999-08-23  
 ; NUMBER OF SEQ ID NOS: 58  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 30  
 ; LENGTH: 627  
 ; TYPE: PRT  
 ; ORGANISM: Abies grandis  
 ; US-09-903-012-30

Query Match 67.7%; Score 2200.5; DB 9; Length 627;  
 Best Local Similarity 68.1%; Pred. No. 1e-201;  
 Matches 427; Conservative 77; Mismatches 106; Indels 17; Gaps 7;

QY	1	MALLSITPLVSRSL-----SSSHEIKALRRRTIPTIGICRPGKSVAHNSNMCLTSVASTD	55
DB	1	MAUVSISPLASKCLRKSLISSIHEHKPPYRTIPNIGMRRRGKSVTPSPMSISLATAADD	60
QY	56	SVORRNVGNSHNLWDDDFIQSLISTPYGADPYRERADRLIGEVKDIMFNFKSLEDGG--	112
DB	61	GVORRIGDYSNINWDDDFIQSL-STPYGEPYQERAEERLIVEVKI-FNSMYLDDGRLMS	118
QY	113	--NDLLQRLVDDVERLIGDRHFKKEIKETALDYVNSYNEKIGGCGRESVVTDLNSTAL	170
DB	119	SPNDLMQRLWIVDSVERLGARHFKNEITSALDYVFRYWEENGIGCGRDSIVTDLNSTAL	178
QY	171	GLRTLRHGVTVSSDVLNVFKDNGQFSSTANTQIEGEIRGVNLFRASIVAPPGKVM	230
DB	179	GFETLRHGVTVSPVVKAFQDQNGQFVCSPG-QTEGEIRSVNLFRASIVAPPGKVM	237
QY	231	EATFTSKYLREALQKIPASSILSLRDLVLEYGWHNLPRLRARNYMDVFGQHTK--N	287
DB	238	EABIFSTRYLKEALQKIPVSA-LSQBIKFVMEYGMHNLPRLEARNYMDVFGQHTK--N	296
QY	288	KNAEKLLELAKLEFNI FHSLORELKHVSRRWKDGSPEMTFCRRHRHVEYYALASCI	347
DB	297	KNAGKKLLELAKLEFNI FNSLOQKELQYLLRMWKESDLPKLTFAHRHVEFFYTLASCI	356
QY	348	EPHSGFRGLGFTKMSHLITVDDMDYVGTVDLELFTATIKWDPSAMECLPEYMGVY	407
DB	357	DPKHSFRGLGFTKMSHLITVDDMDYVGTVDLELFTATIKWDPSAMECLPEYMGVY	416
QY	408	MMVYHTVNEARVAEKAQGRDTLNYARQAEACFDYSYMOEAKWATGYLPTFPEYLENGK	467
DB	477	VSSAYRATLQPTLTNLAWLPDYILKGIDFPSPFNDLASSFLRLGRDTRCYKADRGE	536
QY	528	ASSISCYMDNPGCLTEEDALNHINFMIRDAIRLNWELLKPNDSVPTTSKKHAFDTSRV	587
DB	537	ASCISCYMDNPGSTEDALNHINAMVNDIILKELNELLRSNDNIPMLAKKHAFDITRAL	596
QY	588	HGGRYRDGYSFANVETKSLVMRTVIE	614
DB	597	HHLYIYRDGFSVANKETKLVMTLE	623

RESULT 11  
 US-09-900-797-30  
 ; Sequence 30, Application US/09900797  
 ; Publication No. US20030087406A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chappell, Joseph  
 ; APPLICANT: No. US20030087406A1, Joseph P.  
 ; APPLICANT: Starks, Courtney M.

```
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Abies grandis
US-09-900-797-30

Query Match      67.7%; Score 2200.5; DB 10; Length 627;
Best Local Similarity 68.1%; Pred. No. 1e-201;
Matches 427; Conservative 77; Mismatches 106; Indels 17; Gaps 7;

Qy 1 MALLSITPLVRSCL-----SSSHEIKALRRTIPTLGICRPGKSVAHNINMCLTSVASTD 55
Db 1 MALVSIPLASKSLRKLSSIIHEHKPYTIPNLGMRGRKSVTPSMSISLATAAPDD 60

Qy 56 SVQRRVGNHYNLWDDDFIQSLISTPYGADPYRERADRLIGEVKDIMFNFKSLEDGG-- 112
Db 61 GVQRRIGYHNSINWDDDFIQSL-STPYGEPYQRAERLIVEVKI-FNSMYLDDGRMLS 118

Qy 113 --NDLLQRLLLVDDVERLIGIDRHFKKEIKTALDYVNSYWNKEGICGRESVVTDLNSTAL 170
Db 119 SFNDLMQRLWTVDSVERLIGIARHFKNEITSALDYVFRYWEENGICGGRDSIVTDLNSTAL 178

Qy 171 GRTLRHGYTVSSDVLNVFKDKNGQFSTANIQTGEIRGVNLPRASLVAFPCKEYMD 230
Db 179 GRTLRHGYTVSPVLKAFQDQNGQFVCSPG-QTEGEIRSVNLRYASLAFPGKGYME 237

Qy 231 EAETFTKYLREALQKIPVSA-LSQEIKFVMEYGHWTNLPRLRNRYMDVFGQHTK--N 287
Db 238 EAEFTSTYLKEALQKIPVSA-LSQEIKFVMEYGHWTNLPRLRNRYMDVFGQHTK--N 296

Qy 288 KNAAEKLELAKLEFNIHSLQERELKHVSRWKDGSPEMTFCRRHVEYYALASCIAP 347
Db 297 DPKHSAFRLGFAKMLHVTVDLDYDTFTGIDELFTSAIKRWNSSBIEHLPEYMKCVY 416

Qy 348 EPOHSGRPLGFTKMSHLITVLDMDYDVGTVDELELFTATIKRWDPSAMECLPEYMKGVY 407
Db 417 MVVETVNELTREAKTQGRNTLNVKAWAYFDSYMEAKWISNGYLPMEFVEYHENGK 476

Qy 468 VSSAHRPCALPILTDIPFPDHILKEVDFFPSKUNDLICIILRLRGDTRCYKADRGEE 527
Db 477 VSSAYRVATLOPILTLNWLDPYILKGIDFFSRFNDLASSFLRLRGDTRCYKADRGEE 536

Qy 528 ASSISCYMKNDPGLTEEDALNHINFMIRDAIRELNWELLKPDNSVPITSKKHAFDISRVW 587
Db 537 ASSISCYMKNDPGLTEEDALNHINAMVNDIIRKELNELLNSNDNIPMLAKKHAFDITRAL 596

Qy 588 HHGYRDRGYSFANVETKSLVMRTVIE 614
Db 597 HHLYIYRDGFSVANKETKKLVWETLLE 623

RESULT 12
US-09-893-820-30
; Sequence 30, Application US/09893820
; Publication No. US20040053386A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Chappell, Joseph
; APPLICANT: NO. US20040053386A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Abies grandis
US-09-893-820-30

Query Match      67.7%; Score 2200.5; DB 12; Length 627;
Best Local Similarity 68.1%; Pred. No. 1e-201;
Matches 427; Conservative 77; Mismatches 106; Indels 17; Gaps 7;

Qy 1 MALLSITPLVRSCL-----SSSHEIKALRRTIPTLGICRPGKSVAHNINMCLTSVASTD 55
Db 1 MALVSIPLASKSLRKLSSIIHEHKPYTIPNLGMRGRKSVTPSMSISLATAAPDD 60

Qy 56 SVQRRVGNHYNLWDDDFIQSLISTPYGADPYRERADRLIGEVKDIMFNFKSLEDGG-- 112
Db 61 GVQRRIGYHNSINWDDDFIQSL-STPYGEPYQRAERLIVEVKI-FNSMYLDDGRMLS 118

Qy 113 --NDLLQRLLLVDDVERLIGIDRHFKKEIKTALDYVNSYWNKEGICGRESVVTDLNSTAL 170
Db 119 SFNDLMQRLWTVDSVERLIGIARHFKNEITSALDYVFRYWEENGICGGRDSIVTDLNSTAL 178

Qy 171 GRTLRHGYTVSSDVLNVFKDKNGQFSTANIQTGEIRGVNLPRASLVAFPCKEYMD 230
Db 179 GRTLRHGYTVSPVLKAFQDQNGQFVCSPG-QTEGEIRSVNLRYASLAFPGKGYME 237

Qy 231 EAETFTKYLREALQKIPVSA-LSQEIKFVMEYGHWTNLPRLRNRYMDVFGQHTK--N 287
Db 238 EAEFTSTYLKEALQKIPVSA-LSQEIKFVMEYGHWTNLPRLRNRYMDVFGQHTK--N 296

Qy 288 KNAAEKLELAKLEFNIHSLQERELKHVSRWKDGSPEMTFCRRHVEYYALASCIAP 347
Db 297 DPKHSAFRLGFAKMLHVTVDLDYDTFTGIDELFTSAIKRWNSSBIEHLPEYMKCVY 416

Qy 348 EPOHSGRPLGFTKMSHLITVLDMDYDVGTVDELELFTATIKRWDPSAMECLPEYMKGVY 407
Db 357 DPKHSAFRLGFAKMLHVTVDLDYDTFTGIDELFTSAIKRWNSSBIEHLPEYMKCVY 416

Qy 408 MMVYHTVNMARVAEKACQGRDTLNYARQAEACFDSYMOEAKWIATGYLPTFEYLENGK 467
Db 417 MVVETVNELTREAKTQGRNTLNVKAWAYFDSYMEAKWISNGYLPMEFVEYHENGK 476

Qy 468 VSSAHRPCALPILTDIPFPDHILKEVDFFPSKUNDLICIILRLRGDTRCYKADRGEE 527
Db 477 VSSAYRVATLOPILTLNWLDPYILKGIDFFSRFNDLASSFLRLRGDTRCYKADRGEE 536

Qy 528 ASSISCYMKNDPGLTEEDALNHINFMIRDAIRELNWELLKPDNSVPITSKKHAFDISRVW 587
Db 537 ASSISCYMKNDPGLTEEDALNHINAMVNDIIRKELNELLNSNDNIPMLAKKHAFDITRAL 596

Qy 588 HHGYRDRGYSFANVETKSLVMRTVIE 614
Db 597 HHLYIYRDGFSVANKETKKLVWETLLE 623
```

```
RESULT 13
US-10-041-007-26
; Sequence 26, Application US/10041007
; Publication No. US20020164736A1
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seiichi P.T.
; APPLICANT: Schepmann, Hala G
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase
; FILE REFERENCE: P02081US1
; CURRENT APPLICATION NUMBER: US/10/041,007
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/259,881
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 26
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Abies grandis
US-10-041-007-26

Query Match      67.7%; Score 2200.5; DB 13; Length 627;
Best Local Similarity 68.1%; Pred. No. 1e-201;
Matches 427; Conservative 77; Mismatches 106; Indels 17; Gaps 7;

QY 1 MALLSITPLVSRSL-----SSSHEIKALRRRTIPTLIGICRPGKSVAHNSNMCLTSTVASTD 55
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MALVSIPLASKSLRSLSSISHEHKPPYRTTIPNLGMRGRGKSVTPSPMSISLATAPDD 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 56 SVORRGVNSHNLWDDDFIOSLSTPYGADPYRERADRLIGEVKDMFNPKSLDGG--- 112
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GVORRGIDVSHNLWDDDFIOSL-STPYGEPYSQERAEERLIVEVKKI-FNSMVLDDGRLMS 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 113 --NDLLQRLLLVDDVERLIGIDRHFKEIKTALDYNSYNEKEGIGCGRESVVTDLNSTAL 170
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 119 SFNDLMQRLWIVDSVERLIGARHFKNEITSALDYVFRYWEENGIGCGRDSIVTDLNSTAL 178
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 171 GLRTLRLHGYTVSSDVLNVFKKNGQFSSSTANTQIEGEIRGVNLFRASIVAPFGSKVMD 230
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 GFTLRLHGYTVSPEVLKAFQDQNGQFVCSPG-QTEGEIRSVNLFRASLIAFPGEKWE 237
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 231 EARTFSTKYLREALOKIPASSILSLERDVLEYGWHTNLPRLRARNYMDVFGQHTK--N 287
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 238 EABIFSTRYLKEALQIPVSA-LSQEIKFVMEYGWHTNLPRLRARNYMDVFGQHTK--N 296
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 288 KNAAEKLLLELAKLEFNIFHSLOERELKHVSRRWKKSGSPMTFCRRHRHVEYYALASCI 347
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 297 KNAGKKLLELAKLEFNIFNSLQKELQYLLRWKESDLPKLTFAHRHVEFFYTLASCI 356
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 348 EPOHSGFRLGFTKWSHLITVLDMDYVFGTVDELELFTATIKRWDPSAMECLPEYMKGVY 407
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 357 DPXHSARLGFAGKWCMLVTLDDIYDTFGTIDELELFTSAIKWNSSEIEHLPEYMKCVY 416
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 408 MMVYHTVNMARVAEKAQGRDITNVARQAWAECFDSYMOEAKWIATGYLPTPEYLENGK 467
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 417 MVVFEVNTLTREAEKQGRNTLNLYRKAWAAYFDSYMEAEKWSINGYLPFMEYHENGK 476
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 468 VSSAHRPCALQPILTLDIPFPDHLKEVDFPSPKINDLICIILRLRGDTRCYKADRARGE 527
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 477 VSSAYRVATLQPILTNLNWLDPYILKGIDFPFRFNDLASSFLRLRGDTRCYKADRARGE 536
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 528 ASSISCYMKDNPGLTEDALNHINFMIRDAIRELNWELLKPNDSVPITSKKHAFTDSRYM 587
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 537 ASSISCYMKDNPGTSEDALNHINAMVNDIIEKLNWELLRSNDNIPMLAKKHAFTDITRAL 596
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 588 HHGYRYRDGYSFANVETKSLVMRTVIE 614
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 597 HHLYIYRDGFSVANKETKLVMTLE 623
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 14
US-10-025-145A-2
; Sequence 2, Application US/10025145A
```

```
; Publication No. US20030175861A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Bohlmann, Joerg
; APPLICANT: Steele, Christopher L.
; APPLICANT: Phillips, Michael A.
; TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)
; FILE REFERENCE: WSUR118414
; CURRENT APPLICATION NUMBER: US/10/025,145A
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 09/360,545
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: PCT/US98/14528
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: US 60/052,249
; PRIOR FILING DATE: 1997-07-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Abies Grandis
US-10-025-145A-2

Query Match      67.7%; Score 2200.5; DB 14; Length 627;
Best Local Similarity 68.1%; Pred. No. 1e-201;
Matches 427; Conservative 77; Mismatches 106; Indels 17; Gaps 7;

QY 1 MALLSITPLVSRSL-----SSSHEIKALRRRTIPTLIGICRPGKSVAHNSNMCLTSTVASTD 55
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MALVSIPLASKSLRSLSSISHEHKPPYRTTIPNLGMRGRGKSVTPSPMSISLATAPDD 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 56 SVORRGVNSHNLWDDDFIOSLSTPYGADPYRERADRLIGEVKDMFNPKSLDGG--- 112
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GVORRGIDVSHNLWDDDFIOSL-STPYGEPYSQERAEERLIVEVKKI-FNSMVLDDGRLMS 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 113 --NDLLQRLLLVDDVERLIGIDRHFKEIKTALDYNSYNEKEGIGCGRESVVTDLNSTAL 170
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 119 SFNDLMQRLWIVDSVERLIGARHFKNEITSALDYVFRYWEENGIGCGRDSIVTDLNSTAL 178
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 171 GLRTLRLHGYTVSSDVLNVFKKNGQFSSSTANTQIEGEIRGVNLFRASIVAPFGSKVMD 230
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 GFTLRLHGYTVSPEVLKAFQDQNGQFVCSPG-QTEGEIRSVNLFRASLIAFPGEKWE 237
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 231 EARTFSTKYLREALOKIPASSILSLERDVLEYGWHTNLPRLRARNYMDVFGQHTK--N 287
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 238 EABIFSTRYLKEALQIPVSA-LSQEIKFVMEYGWHTNLPRLRARNYMDVFGQHTK--N 296
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 288 KNAAEKLLLELAKLEFNIFHSLOERELKHVSRRWKKSGSPMTFCRRHRHVEYYALASCI 347
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 297 KNAGKKLLELAKLEFNIFNSLQKELQYLLRWKESDLPKLTFAHRHVEFFYTLASCI 356
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 348 EPOHSGFRLGFTKWSHLITVLDMDYVFGTVDELELFTATIKRWDPSAMECLPEYMKGVY 407
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 357 DPXHSARLGFAGKWCMLVTLDDIYDTFGTIDELELFTSAIKWNSSEIEHLPEYMKCVY 416
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 408 MMVYHTVNMARVAEKAQGRDITNVARQAWAECFDSYMOEAKWIATGYLPTPEYLENGK 467
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 417 MVVFEVNTLTREAEKQGRNTLNLYRKAWAAYFDSYMEAEKWSINGYLPFMEYHENGK 476
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 468 VSSAHRPCALQPILTLDIPFPDHLKEVDFPSPKINDLICIILRLRGDTRCYKADRARGE 527
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 477 VSSAYRVATLQPILTNLNWLDPYILKGIDFPFRFNDLASSFLRLRGDTRCYKADRARGE 536
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 528 ASSISCYMKDNPGLTEDALNHINFMIRDAIRELNWELLKPNDSVPITSKKHAFTDSRYM 587
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 537 ASSISCYMKDNPGTSEDALNHINAMVNDIIEKLNWELLRSNDNIPMLAKKHAFTDITRAL 596
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 588 HHGYRYRDGYSFANVETKSLVMRTVIE 614
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 597 HHLYIYRDGFSVANKETKLVMTLE 623
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Db 597 HHLVYRDGFSVANKETKLVMTLLE 623

Search completed: July 7, 2004, 17:45:19

Job time : 53 secs

```
RESULT 15
US-10-025-145A-32
; Sequence 32, Application US/10025145A
; Publication No. US20030175861A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Bohmann, Joerg
; APPLICANT: Steele, Christopher L.
; APPLICANT: Phillips, Michael A.
; TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)
; FILE REFERENCE: WSU118414
; CURRENT APPLICATION NUMBER: US/10/025,145A
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 09/360,545
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: PCT/US98/14528
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: US 60/052,249
; PRIOR FILING DATE: 1997-07-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Abies Grandis
US-10-025-145A-32

Query Match 67.6%; Score 2197.5; DB 14; Length 627;
Best Local Similarity 68.1%; Pred. No. 2e-201;
Matches 427; Conservative 77; Mismatches 106; Indels 17; Gaps 7;

Qy 1 MALLSITPLVRSCL-----SSSHEIKALRRTIPTPLGICRPGKSVAHNINCLTSVASTD 55
Db 1 MALVSIPLASKSLRKSLISSIHEHKPPYRTIPNLGMRRRGKSVTPSMSISLATAAPDD 60

Qy 56 SVQVRGVNHYNLDDDDFIQSLSTPYCAPDYRERADRLIGEVKDJMENPKSLBDGG--- 112
Db 61 GVQRRIGDYHNSIWDDDFIQSL-STHYGEPSYQERAERLIVEVKKI-FNSMYLDDGR LMS 118

Qy 113 --NDLLQRLLLVDVERIGIDIRHFKKEIKTALDYVNSYWNKEGICGGRSVVTDLNSTAL 170
Db 119 SFNDLMQELWIVDSVERLGIARHFKNEITSALDYVFRYWEENGICGGRDSIVTDLNSTAL 178

Qy 171 GLRILRLHGYTVSSDLNVFKDKGQFSSTANIQIEGIRGVNLNFRASLVAFPGEKYMD 230
Db 179 GFRTILRHGYTVSPVLKAFQDQNGQFVCSPG-QTEGEIRSVNLNLYRASLIAFPGEKYME 237

Qy 231 EAETFTKYLREALOKIPASSITLSLEIRDVLEYGWHNTLPLEARNYMDVFGQHTK---N 287
Db 238 EAEIFSTRYLKEALQKIPVSA-LSOEIRFVMEYGWHNTLPLEARNYIDTLEKOTSAWLN 296

Qy 288 KNAAEKLLLEAKLEFNIPHSIQERELKHVSRWWDKSGSPENTFCRHRHVEYYVALASCI AF 347
Db 297 KNAGKULLEAKLEFNIPNSIQKELQYLLRWKESDLPKLTFAHRHVEFYTLASCI AI 356

Qy 348 EPQHSGRFLGTGKSHLITVLDMDYVGTGTVDELELFTATIKRWDPPSAMECLPEYMKGVY 407
Db 357 DPKHSAPFLGFAKMLHVTVLDLDIYDTFTGTTDELELFTSAIKRWNSSEIEHLPEYMKCVY 416

Qy 408 MMVYHTVNMARVAEKAQGRDNLNVAQAEACFDSYMQEAKWIATGYLPTFEEYLENGK 467
Db 417 MVVFETVNLTRAEAKTQGRNTLNVKRAWEAYFDSYMEEAKWINSGLYPTFEEYHENGK 476

Qy 468 VSSAHRPCALQPIFLDIPFPDHLKVEYDFPSKLNLDLICIILRLRGDTRCYKADRGEE 527
Db 477 VSSAVRVATLQPIFLTNALWLPDYILKGIDFDSRFENDLASSFLRLRGDTRCYKADRGEE 536

Qy 528 ASSISCYKNDPGLTEEDALNHINMIRDAIRLWELLKPDNSVPTSKXKHAFDISRW 587
Db 537 ASCISCYKNDPGLTEEDALNHINAMVNDIIKELWELLIRSDNIPMLAKKHAFDITRAL 596

Qy 588 HHGYRYRDGYGFANVETKSLVMRTVIE 614
Db 588 HHGYRYRDGYGFANVETKSLVMRTVIE 614
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 8, 2004, 12:35:55 ; Search time 6774 Seconds  
(without alignments)  
3954.236 Million cell updates/sec

Title: US-10-025-145A-65

Perfect score: 3251

Sequence: 1 MALLSITPLVSRSLSSSHE.....FANVETKSLVNRVTIEVPVL 618

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Deiop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US10025145/runat\_07072004\_160149\_27936/app.query.fasta\_1.775  
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORES=pcct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US10025145 @cgn 1.1 4198 @runat\_07072004\_160149\_27936 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPOP=10 -XGAPEXT=0.5 -XGAPOP=6  
-XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_rscs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3251	100.0	2013	6	AR222136	Sequence
2	3251	100.0	2013	8	AGU87910	U87910 Abies grand
3	2456.5	75.6	1958	8	AY237645	AY237645 Picea sit
4	2453	75.5	2018	6	BD227661	BD227661 Syntheses
5	2453	75.5	2018	6	AR222097	AR222097 Sequence
6	2453	75.5	2018	6	AR240696	AR240696 Sequence
7	2453	75.5	2018	6	AR266971	AR266971 Sequence
8	2453	75.5	2018	6	AR316320	AR316320 Sequence
9	2453	75.5	2018	6	AR338463	AR338463 Sequence
10	2453	75.5	2018	6	AR429869	AR429869 Sequence
11	2453	75.5	2018	8	AGU87909	U87909 Abies grand
12	2328.5	71.6	2082	8	AF543527	AF543527 Pinus tae
13	2266.5	69.7	2162	8	AF543529	AF543529 Pinus tae
14	2264	69.6	2186	6	AR222137	AR222137 Sequence
15	2264	69.6	2186	8	AF139205	AF139205 Abies gra
16	2237	68.8	2198	8	AF369918	AF369918 Picea abi
17	2205.5	67.8	2150	8	AF369919	AF369919 Picea abi
18	2200.5	67.7	2196	6	BD227666	BD227666 Syntheses
19	2200.5	67.7	2196	6	AR222096	AR222096 Sequence
20	2200.5	67.7	2196	6	AR240701	AR240701 Sequence
21	2200.5	67.7	2196	6	AR266976	AR266976 Sequence
22	2200.5	67.7	2196	6	AR316325	AR316325 Sequence
23	2200.5	67.7	2196	6	AR338468	AR338468 Sequence
24	2200.5	67.7	2196	6	AR429874	AR429874 Sequence
25	2200.5	67.7	2196	8	AGU87908	U87908 Abies grand
26	2197.5	67.6	2205	6	AR222116	AR222116 Sequence
27	2163.5	66.5	2186	8	AF461460	AF461460 Picea abi
28	2151	66.2	2100	8	AF543530	AF543530 Pinus tae
29	2137	65.7	1890	6	AR222146	AR222146 Sequence
30	2137	65.7	1893	8	AF139206	AF139206 Abies gra
31	2084.5	64.1	2429	6	AR222138	AR222138 Sequence
32	2084.5	64.1	2429	8	AF139207	AF139207 Abies gra
33	2046.5	62.9	3306	8	AF326517	AF326517 Syntheses
34	2031.5	62.5	2089	6	BD227677	BD227677 Syntheses
35	2031.5	62.5	2089	6	AR222098	AR222098 Sequence
36	2031.5	62.5	2089	6	AR240712	AR240712 Sequence
37	2031.5	62.5	2089	6	AR266987	AR266987 Sequence
38	2031.5	62.5	2089	6	AR316336	AR316336 Sequence
39	2031.5	62.5	2089	6	AR338479	AR338479 Sequence
40	2031.5	62.5	2089	6	AR429885	AR429885 Sequence
41	2031.5	62.5	2089	8	AF006193	AF006193 Abies gra
42	1985.5	61.1	1960	8	AF543531	AF543531 Pinus tae
43	1921.5	59.1	2084	8	AF543528	AF543528 Pinus tae
44	1699	52.3	3116	8	AF326518	AF326518 Abies gra
45	1617.5	49.8	1513	6	BD272958	BD272958 Materials

#### ALIGNMENTS

RESULT 1

AR222136  
LOCUS AR222136 2013 bp DNA linear PAT 26-SEP-2002  
DEFINITION Sequence 64 from patent US 6429014.  
ACCESSION AR222136  
VERSION AR222136.1 GI:23329510  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1. (bases 1 to 2013)  
Unclassified.  
AUTHORS Steele,C.L., Bohlmann, J. and Croteau,R.B.  
TITLE Monoterpene synthases from grand fir (*Abies grandis*)  
JOURNAL Patent: US 6429014-A 64 06-AUG-2002;  
FEATURES  
Location/Qualifiers  
1..2013  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,03e-255 Length: 2013  
Score: 3251.00 Matches: 618  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-10-025-145A-65 (1-618) x AR222136 (1-2013)  
QY 1 MetAlaLeuSerIleThrProLeuValSerArgSerCysLeuSerSerHisGlu 20  
Db 36 ATGGCTCTCTCTTCTATTACTCGCTGGTTTCAGGTCGTCCTCAGTCTCTCTCATGAG 95  
QY 21 IleLysAlaLeuArgThrIleProThrLeuGlyIleCysArgProGlyLysSerVal 40  
Db 96 ATTAAGGCTCTCCGTAGAACAAATCCCACTCTTGGAACTCTGGAACTCTGGAGCGGGGAAATCCGTC 155  
QY 41 AlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAspSerValGlnArgArg 60  
Db 156 GCGCATTCCTAATAACATGTTGTTGACAAGCGTCGATCTACTGATTCGTACAGAGACGC 215  
QY 61 ValGlyAsnIleHisSerAsnLeuThrAspAspPheIleGlnSerIleSerThr 80  
Db 216 GTGGCAACTATCATTTCCAACTCTGGGACGATGATTTTCATACAGTCTCTGATCTCAAGC 275  
QY 81 ProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIleGlyGluValLysAsp 100  
Db 276 CCTTATGGAGCCTGATTACCGGGAACGTCTGACAGACTTATTGGGGAAGTAAAGGAT 335  
QY 101 IleMetPheAsnPhelyssSerLeuGluAspGlyGlyAsnAspLeuLeuGlnArgLeuLeu 120  
Db 336 ATAATGTTCAATTTCAAGTCGCTGGGAAGATGGAGCAATGATCTCTTCAACGACATTTTG 395  
QY 121 LeuValAspValGluArgLeuGlyIleAspArgHisPhelyssGlyLysGluIleLysThr 140  
Db 396 CTGGTCGATGACGTTGAACGTTTGGGAATCGACAGGCATTTCAAAAAAGAGATAAAACG 455  
QY 141 AlaLeuAspTyrValAsnSerTyrTrpAsnGluLysGlyIleGlyCysGlyArgGluSer 160  
Db 456 GCACCTGANTATTGTTACAGTTATTGGAAACGAAAAGGCATTTGGATGTGGAGGGAGAGT 515  
QY 161 ValValThrAspLeuAsnSerThrAlaLeuGlyLeuArgThrLeuArgLeuHisGlyTyr 180  
Db 516 GTTGTGACGTGACCTCAACTCAACCGCTTTGGGCTTCGAACTCTCCGACTACAGCGATAC 575  
QY 181 ThrValSerSerAspValLeuAnValPheLysAspLysAsnGlyGlnPheSerSerThr 200  
Db 576 ACTGTGCTCTCAGATGTTTGAACGTTTTTAAAGACAAAATGGGCAATTTTCTCCACT 635  
QY 201 AlaAsnIleGlnIleGluGlyIleArgGlyValLeuAsnLeuPheArgAlaSerLeu 220  
Db 636 GCCAATATTACAGATAGAGGGAGAGATTAGAGGGGTTCTCAATTATTTCAGGGGCTCCCTC 695

QY 221 ValAlaPheProGlyGluLysValMetAspGluAlaGluThrPheSerThrLysTyrLeu 240  
Db 696 GTCGCCTTTCCCGCGGAGAAAGTTATGGATGAAGCTGAAACATTTCTCTCAAAATATTTA 755  
QY 241 ArgGluAlaLeuGlnLysIleProAlaSerSerIleLeuSerLeuGluIleArgAspVal 260  
Db 756 AGAGAAGCCCTGCAAAAGATTCCGGCATCCAGTATATTCTTCACTAGAGATACGGGACGTT 815  
QY 261 LeuGluTyrGlyTyrIleThrAsnLeuProArgLeuGluAlaArgAsnTyrMetAspVal 280  
Db 816 CTGGATATGGTTGGCACACCAATTTGCCACGCTTGGAAAGCAAGAAATTCATCGACGTC 875  
QY 281 PheGlyGlnHisThrLysAsnLysAsnAlaAlaGluLysLeuGluLeuAlaLysLeu 300  
Db 876 TTTGGACAGCACACTAAAAATAAGAACGCGCGCGGAAACTTTTAGAACTTGCAAAATTG 935  
QY 301 GluPheAsnIlePheHisSerLeuGlnArgGluLeuLysHisValSerArgTyrTrp 320  
Db 936 GAATTCATATATTTTCACTCTTACAAGAGAGAGAGTTAAAAACATGTTTCCCGATGGTGG 995  
QY 321 LysAspSerGlySerProGluMetThrPheCysArgHisArgHisValGluTyrTyrAla 340  
Db 996 AAGACTCGGGTTCTCTGAGATGACCTTCTGTGCACATCGTCACGTGGAATACTACGCT 1055  
QY 341 LeuAlaSerCysIleAlaPheGluProGlnHisSerGlyPheArgLeuGlyPheThrLys 360  
Db 1056 TTGGCTTCTCGCATTCGGTTTCGAGCTCAACATTTCTGGATTCTGAGCTCGGCTTACCAAG 1115  
QY 361 MetSerHisLeuIleThrValLeuAspAspMetTyrAspValPheGlyThrValAspGlu 380  
Db 1116 ATGTCTCATCTTATACGGTTCTTACGACATGTACGACGTCTTCGGCAGTAGACGAG 1175  
QY 381 LeuGluLeuPheThrAlaThrIleLysArgTyrAspProSerAlaMetGluCysLeuPro 400  
Db 1176 CTGGAACCTCTTACACGACCAATTAAGAGATGGATCCGTCCGCGATGGAATGCCTTCCA 1235  
QY 401 GluTyrMetLysGlyValTyrMetMetValTyrHisThrValAsnGluMetAlaArgVal 420  
Db 1236 GAATATATGAAAGGAGTGTACATGATGTTTATACACCGTAAATGAAATGGCTCGAGTG 1295  
QY 421 AlaGluLysAlaGluGlyArgAspThrLeuAsnTyrAlaArgGlnAlaCys 440  
Db 1296 GCAGAGAAGGCTCAAGCGCGACAGCCTCAACTATGCAAGACAGCGCTTGGGAGGGGTGT 1355  
QY 441 PheAspSerTyrMetGlnGluAlaLysTyrIleAlaThrGlyTyrLeuProThrPheGlu 460  
Db 1356 TTTGATTCGTATATGACGAAGCAAGTGCATCGCCATGGTTATCTGCCCGCTTTGAG 1415  
QY 461 GluTyrLeuGluAsnGlyLysValSerSerAlaHisArgProCysAlaLeuGlnProIle 480  
Db 1416 GAGTACTTCGAGAACGGGAAAGTTAGTCTCTGCTCATCGCCCATCGCACTGCAACCCATT 1475  
QY 481 LeuThrLeuAspIleProPheProAspHisIleLeuLysGluValAspPheProSerLys 500  
Db 1476 CTGACGTTGGATCCCTTTCTGATCATCATCTTCAAGAAAGTTGACTTCCATCCGAG 1535  
QY 501 LeuAsnAspLeuIleCysIleLeuArgLeuArgGlyAspThrArgCysTyrLysAla 520  
Db 1536 CTCATGACTTCATATGATATCATCTTCGATTAAAGAGGTGATACACGGTGTCTACAAGCA 1595  
QY 521 AspArgAlaArgGlyGluAlaSerSerIleSerCysTyrMetLysAspAsnProGly 540  
Db 1596 GACAGGGCCGCTGGGAGAAAGCTTCGTCTATATCATATGTTATATGAAGACAAATCTCGA 1655  
QY 541 LeuThrGluGluAspAlaLeuAsnHisIleAsnPheMetIleArgAspAlaIleArgGlu 560  
Db 1656 TTAACGGAAGAGAGCTCTGAATCATATCAACTTCATGATCAGGAGCCCAATTCAGAGAA 1715  
QY 561 LeuAsnTrpGluLeuLeuLysProAspAsnSerValProIleThrSerLysLysHisAla 580  
Db 1716 TTAATTTGGAGCTTCTAAAGCCAGACACAGTGTTCCTCATCTTCCAAAGAAACACGCA 1775  
QY 581 PheAspIleSerArgValTrpHisGlyTyrArgTyrArgAspGlyTyrSerPheAla 600

```
|||||
Db 1776 TTTGACATAACAGAGTTTGGCATCACGTTACAGATACCGAGATGGCTACAGCTTGGC 1835
Qy 601 AsnValGluThrLysSerLeuValMetArgThrValIleGluProValProLeu 618
1836 AACCTTGAACAAGAGTTTGGTATGAGACCGTCATGACCTGTGCGCTTTG 1889

RESULT 2
AGU87910 2013 bp mRNA linear PLN 10-AUG-2001
DEFINITION Abies grandis (-)-camphene synthase (AG6.5) mRNA, complete cds.
ACCESSION U87910
VERSION U87910.1 GI:2411484
KEYWORDS :
SOURCE Abies grandis
ORGANISM Abies grandis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
REFERENCE 1 (bases 1 to 2013)
Bohlmann, J., Steele, C.L. and Croteau, R.
Monoterpene synthases from grand fir (Abies grandis). cDNA
isolation, characterization, and functional expression of myrcene
synthase, (-)-(4S)-limonene synthase, and (-)-(1S,5S)-pinene
synthase
J. Biol. Chem. 272 (35), 21784-21792 (1997)
JOURNAL MEDLINE 97413772
PUBMED 9268308
REFERENCE 2 (bases 1 to 2013)
Bohlmann, J., Phillips, M., Ramachandiran, V., Katoh, S. and Croteau, R.
cDNA cloning, characterization, and functional expression of four
new monoterpene synthase members of the Tpsd gene family from grand
fir (Abies grandis)
Arch. Biochem. Biophys. 368 (2), 232-243 (1999)
JOURNAL MEDLINE 99373092
PUBMED 1041373
REFERENCE 3 (bases 1 to 2013)
Bohlmann, J., Steele, C.L. and Croteau, R.
Direct Submission
Washing State University, Clark Hall, Pullman, WA 99164-6340,
USA
FEATURES
source 1..2013
/organism="Abies grandis"
/mol_type="mRNA"
/db_xref="taxon:46611"
gene 1..2013
/gene="AG6.5"
CDS 36..1892
/gene="AG6.5"
/note="terpene synthase; monoterpene synthase; terpene
cyclase"
/codon_start=1
/product="(-)-camphene synthase"
/protein_id="AAB70707.1"
/db_xref="GI:2411485"
/translacion="MALLSITPLVSRSLSSSHBIKALRRITPTLGI CRPGKSVASHI
NMCLTSVASTDSVQRRVYNSHNLWDDDFIQSLSTPYGADPYRERADRLIGEVKDIM
FNFKSLDGGNDLQRLLVDDVERLGDIDRHFKEIKTALDYVNSYWNKXGICGGRS
VVDNLSTALGLRTLRLHGYTVSSDLNVKPKNGQSPSTANIOIEGIRGVNLNFA
SLVAFPGKVDENAEFTSKVLRALQKIPASSILSLERDVLBYGHHTNLPLEARN
YMDVFGQHTKNNAEKLELAKLEFNIFHSLOBERLKHVSRRWKKDSGSPMTFCRRH
HVEYVALASCIATFQPHSGFRLGTRKMSHLITVLDDMDYDVFQVDELELEFTATIKRW
PSAMECLPEYMKGYMVTNEMARVAERKAQGRDITNVARQAEACFDYSMOEAKW
IATGVLPTFERYLNGKVSSAHRKALQPLITLDIPDPDHLKEVDYPSKLDLICI
LRLRGDRRCYKADRGAEASSISCYMKDNPGLTEEDALNHINFEIDALRELNWELL
KPDNSVITPSKKHAFDLSRVWHHGYRYRDGYSFANVETKSLVMVTIEVPVL"
```

## ALIGNMENT SCORES:

```
Pred. No.: 2,03e-255 Length: 2013
Score: 3251.00 Matches: 618
Percent Similarity: 100.00% Conservative: 0
```

```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0
US-10-025-145a-65 (1-618) x AGU87910 (1-2013)
Qy 1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeuSerSerHisGlu 20
Db 36 ATGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 95
Qy 21 IleLysAlaLeuArgArgThrIleProThrLeuGlyIleCysArgProGlyVal 40
Db 96 ATTAAAGGCTCTCCGTAGAACCAATCCCAACTCTTGAATCTCGAGCCGGGAAATCCGTC 155
Qy 41 AlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAspSerValGlnArgArg 60
Db 156 GCGCATTCCTATAAATCATGTGTTGCACAGCGTCGCATCTACTGATCTGTACAGAGCGC 215
Qy 61 ValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIleGlnSerLeuIleSerThr 80
Db 216 GTGGGCAACTATCATTCACACCTGTGGAGCAGATGATTTTCATACAGTCTCTGATCTCAACG 275
Qy 81 ProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIleGlyGluValIleAsp 100
Db 276 CCTATGAGCAGCCTGATTTACCGGGAACGTGCTGACAGACTTATTGGGGAAGTAAAGAT 335
Qy 101 IleMetPheAsnPheLysSerLeuGluAspGlyGlyAsnAspLeuLeuGlnArgLeuLeu 120
Db 336 ATAATGTTCAATTTCAAGTCGTCGGAAGTGGAGCAATGATCTCTTCAACGACTTTTG 395
Qy 121 LeuValAspAspValGluArgLeuGlyIleAspArgHisPheLysLysGluIleLysThr 140
Db 396 CTGGTCGATGACGTGTGAACGTTTGGGAATCGACAGGCAATTTCAAAAAGAGATAAAAACG 455
Qy 141 AlaLeuAspTyrValAsnSerTyrTrpAsnGluLysGlyIleGlyCysGlyArgGluSer 160
Db 456 GCATCTCGATATGTTAATACAGTTATTGGAAACGAAAGCAATGGATGTGGAGGAGAGT 515
Qy 161 ValValThrAspLeuAsnSerThrAlaLeuGlyLeuArgThrLeuArgLeuHisGlyTyr 180
Db 516 GTTGTGACTGACCTCAACTCAACCGCTTGGGGCTTCGAACCTCTCCGACTACACGGATAC 575
Qy 181 ThrValSerSerAspValLeuAsnValPheLysAspLysAsnGlyGlnPheSerSerThr 200
Db 576 ACTGTGCTCTCAGATGTTTGAACGTTTTTAAAGACAAATAATGGCAATTTTCTCCACT 635
Qy 201 AlaAsnIleGlnIleGluGlyValArgGlyValLeuAsnLeuPheArgAlaSerLeu 220
Db 636 GCCAATATTGATAGTAGGAGAGATTTAGAGCGCTTCTCAATTTATTTCAGGGCCCTCCCTC 695
Qy 221 ValAlaPheProGlyGlyLysValMetAspGluAlaGluThrPheSerThrLysTyrIleu 240
Db 696 GTCGCTTTCCCGCGAGAAAAGTTATGGATGAAGCTGAAACATTTCTCTACAAAATATTTA 755
Qy 241 ArgGluAlaLeuGlnLysIleProAlaSerSerIleLeuSerLeuGluIleArgAspVal 260
Db 756 AGAAGAACCCCTGCAAAAAGATTTCCGCGCATCCAGTATACTTTTCACTAGAGATACGGGACGTT 815
Qy 261 LeuGluTyrGlyTyrHisThrAsnLeuProArgLeuGluAlaArgAsnTyrMetAspVal 280
Db 816 CTGGAATATGGTTGGCACCAATTTGCCAGCTTTGGAAGCAAGAAATTTACATGAGCGTC 875
Qy 281 PheGlyGlnHisThrLysAsnLysAsnAlaGluLysLeuLeuGluLeuAlaLysLeu 300
Db 876 TTTGACAGACACACTAAAAATAAGAACCGCCCGAGAAACTTTTAGAACTTGCAAAATTG 935
Qy 301 GluPheAsnIlePheHisSerLeuGlnGluArgGluLeuLysHisValSerArgTyrTrp 320
Db 936 GAATTCATATATTTCATCTCTCTTACAAGAGAGAGAGTAAACAATGTTTCCCGATGTGTGG 995
Qy 321 LysAspSerGlySerProGluMetThrPheCysArgHisArgHisValGluTyrTyrAla 340
Db 996 AAAGACTCGGGTCTCTCTGAGATGACCTTCTGTGACATGTCACGTGGGAATATCTACGCT 1055
```

Qy	341	LeuAlaSerCysIleAlaIleAlaPheGluProGlnHisSerGlyPheArgLeuGlyPheThrLys	360
Db	1056	TTGGCTTCCTCGCATTCGGTTTCGAGCCTCAACATTCCTGGATTCAGACTCGGCTTACCAAG	1115
Qy	361	MetSerHisIleuIleThrValLeuAspAspMetTyrAspValPheGlyThrValAspGlu	380
Db	1116	ATGTCCTCATCTTATCAGCGGTCTTTGACGATGTACAGCTCTTCGGCAGAGTAGACGAG	1175
Qy	381	LeuGluLeuPheThrAlaThrIleLysArgTTPAspProSerAlaMetGluCysLeuPro	400
Db	1176	CTGGAACTCTCACAGCGACAAATTAAAGAGATGGATCCGTCCGCGATGGATGCCCTTCCA	1235
Qy	401	GluTyrMetLysGlyValTyrMetMetValTyrHisThrValAsnGluMetAlaArgVal	420
Db	1236	GAATATATGAAGAGAGTGCATCATGATGGTTTATCACACCGTTAAATGATGAAATGGCTCGAGTG	1295
Qy	421	AlaGluLysAlaGlnGlnIleArgAspThrLeuAsnTyrAlaArgGlnAlaTTPGluAlaCys	440
Db	1296	GCAGAGAGAGCTCAAGCCCGAGACACGCTCACTATGCAAGACAGGCTTGGGAGGCGTGT	1355
Qy	441	PheAspSerTyrMetGlnGlnAlaLysTTPileAlaThrGlyTyrLeuProThrPheGlu	460
Db	1356	TTTGATTTCGTATATGCAAGAGCAAAAGTGGATCGCCACTGTTATCTTCGCCACAGTTTGAG	1415
Qy	461	GluTyrLeuGluAsnGlnLysValSerSerAlaHisArgProCysAlaLeuGlnProIle	480
Db	1416	GAGTACTTTGGAAACGGGAAGTTAGTCTCTGCTCATCGCCATCGGCACATGCAACCCATT	1475
Qy	481	LeuThrLeuAspIleProPheProAspHisIleLeuLysGluValAspPheProSerLys	500
Db	1476	CTGACGTTGGACATCCCTTTCCTGTGATCATCATCTCAAGGAAGTTGACTTCCCATCGAAG	1535
Qy	501	LeuAsnAspLeuIleCysIleIleLeuArgLeuArgGlyAspThrArgCysTyrLysAla	520
Db	1536	CTCAATGACATTGATATGATATCATCTCTTCGATTAAAGAGTGATACACGGTGCTCAAGGCCA	1595
Qy	521	AspArgAlaArgGlyGluGluAlaSerSerIleSerCysTyrMetLysAspAsnProGly	540
Db	1596	GACAGGGCCCGTGGAGAGAAAGCTTCGTCTATATCATGTTATATGAAGACAACTCCTGGA	1655
Qy	541	LeuThrGluGluAspAlaLeuAsnHisIleAsnPheMetIleArgAspAlaIleArgGlu	560
Db	1656	TTAACGGGAAGAAGATGCTCTGAAATCATATCAACTTCATGATCAGGGACGCAATCAGAGAA	1715
Qy	561	LeuAsnTrpGluLeuLeuLysProAspAsnSerValProIleThrSerLysLysHisAla	580
Db	1716	TTAAATTTGGAGACTCTTAAGCCAGACAAACAGTGTCCCATCATCTTCCAGAAACACGCA	1775
Qy	581	PheAspIleSerArgValTPHisHisGlyTyrArgTyrArgAspGlyTyrSerPheAla	600
Db	1776	TTTGACATAAGCAGAGTTTGGCATCCGGTTACAGATACCGAGATGGCTACAGCTTTGCC	1835
Qy	601	AsnValGluThrLysSerLeuValMetArgThrValIleGluProValProLeu	618
Db	1836	AACGTTGAACAAAGAGTTTGGTATGAGAACCGTCAATTGAACCTGTGCGCTTG	1899
RESULT 3			
AY237645			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Picea sitchensis (Sitka spruce)			
Picea sitchensis			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.			
1 (bases 1 to 1958)			
Bryn McKay, S.A., Hunter, W.L., Godard, K.A., Wang, S.X., Martin, D.M.,			
Bohlmann, J. and Plant, A.L.			
Insect Attack and Wounding Induce Traumatic Resin Duct Development			
and Gene Expression of (-)-Pinene Synthase in Sitka Spruce			
TITLE			
AY237645 1958 bp mRNA linear PLN 13-SEP-2003			
Picea sitchensis pinene synthase mRNA, complete cds.			
AY237645.1 GI:34582666			

Plant Physiol. 133 (1), 368-378 (2003)  
12970502  
2 (bases 1 to 1958)  
Byun McKay,A.S., Hunter,W.L., Goddard,K.A., Wang,S.X., Martin,D.,  
Bohlmann,J. and Plant,A.L.  
Direct Submission  
Submitted (17-FEB-2003) Biology, Simon Fraser University, 8888  
University Drive, Burnaby, BC V5A 1S6, Canada  
Location/Qualifiers  
1. 1958  
/organism="Picea sitchensis"  
/mol\_type="mRNA"  
/db\_xref="taxon:3332"  
/clone="19-2-1/12-3"  
75. 1958  
/note="terpene synthase; monoterpene synthase; terpene  
cyclase; (-)-pinene synthase"  
/codon\_start=1  
/product="pinene synthase"  
/protein\_id="AAP72020.1"  
/db\_xref="GI:34582667"  
/translation="MALVSVAPMASRSLHKSLSSSAHLEKLTICRTPIGLGMRRRKKS  
ATKSMGSLTTSYDDGQRMRGDFFSNLWDDFIQSLSTSGEYSPYRAERAEILGEV  
KMFNMSGSDGELISPHNDILQRVMWDSVERLGEHFRFNEIKGALDYVYSGWSEK  
GIGCGSVVADLNTGALFPTLRHLGYAVSADLVKLDQNGQACFSQSPTEERIS  
VLNLRYASLTFAPGKMEAEATFSAKYLEEESLOKLSVSSLOEIRDLVEYGWHTVLP  
RMEARHIDVFGDTONSKSCINTEKLLLEKLEFNI FHSLOKRELEYLVRWKQDSGS  
PQMTFCRRHVEADTILASCFAPBQSGFRGLGFKAKCHITILDDMYDTFGTVDELEL  
FTWTFRRHDSADCPKSGMGVYLILYDVTNETSREAKQQRDTLDYVARRADELLYL  
DSYMQAKMTATGYPTEAYENBGKVSQGHRTSALQPIITWMDIPFPFHLLKEVDFFS  
KLNDLASAILRLGDRCTCYKADRGAEASSISCYMKDNPFGATEADALDHINAMISDV  
IRGLNELLNPNSSVPISSKKHVPDISRAPHYGYKYVRDGYSVANIETKSLVRRVTVIDP  
VTTL"

## ORIGIN

```

Alignment Scores:
Pred. NO.:      1.17e-190      Length:      1958
Score:          2456.50         Matches:     473
Percent Similarity: 84.79%      Conservative: 62
Best Local Similarity: 74.96%   Mismatches:  79
Query Match:     75.56%        Indels:     17
DB:              8             Gaps:        7

US-10-025-145A-65 (1-618) x AY237645 (1-1958)

Qy  1  MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15
Db  75  ATGGCTCTGGTTCTGTGGCCCTATGGCTTCAGATCATGCCTGCACAATCGTTGAGC 134
Qy  16  SerSerSerHisGluIleLysAlaLeuArgArgThrIleProThrLeuGlyIleCysArg 35
Db  135  AGTTCTGCTCAGGAGCTTAGACGATCTGTAGAACAAATCCCAACTCTTGGAAATGCTAGG 194
Qy  36  ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55
Db  195  CGAGGAAATCTCGGACGCCCTTCTATGAGCATGAGTTTGACCACCACTGATCTGATGAT 254
Qy  56  SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuThrAspAspAspPheIleGln 75
Db  255  GGTGTACAAAGACGGATGGCGAATTCATTCCAACCTATGGAACCGATGATTTTCATACAG 314
Qy  76  SerLeuIleSerThrProTyrGlyAlaProCaspTyrArgGluArgAlaAspArgLeuIle 95
Db  315  TCTCTC---TCAAGCTCTTAATGGGAACTTCTTACCGGGAAACGAGCTGAGACTGTATT 371
Qy  96  GlyGluValLysAspIleMetPheAsnPheLysSerLeuGluAspGly----- 111
Db  372  GGGGAAGTAAAGAAG--ATGTTCAATTCAATGTCATCGGAACGAGAGAGTTAATCAGT 428
Qy  112  ---GlyAsnAspLeuLeuGlnArgLeuLeuLeuValAspAspValGluArgLeuGlyIle 130
Db  429  CCCACAAATGATCTCTCAATCAACGCGTTTGGATGGTTGATGTTGAACGCTTTAGGGATA 488

```

US-10-025-145A-65 (1-618) X AY237645 (1-1958)

Qy	1	MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu-----	15
Db	75	ATGGCTCTGGTCTTCTTGGCGCCTATGGCTTCCAGATCATGCGTGCACAATCGTTGAGC	134
Qy	16	SerSerSerHisGluIleLeuAlaLeuArgArgThrIleProThrLeuGlyIleCysArg	35
Db	135	AGTTCTGCTCACGAGCTTAAGACGATCTGTAGAACAAATCCCACTCTTGGAAATGCTTAGG	194
Qy	36	ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp	55
Db	195	CGAGAAATCTCGACGCCTCTATGAGCATGAGTTTGACCACTGATCTCTGATGAT	254
Qy	56	SerValGlnArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIleGln	75
Db	255	GGTGTCAAAGACGGATGGGCGAATTCATTCCAACCTATGGAACGATGATTTTCATACAG	314
Qy	76	SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle	95
Db	315	TCTCTC---TCAACGCTTTATGGGAACTTCTTTACGGGAACGAGCTGAGACTGATT	371
Qy	96	GlyGluValLysAspIleMetPheAsnPheLysSerLeuGluAspGly-----	111
Db	372	GGGGAAGTAAAGAG---ATGTTTCAATTCAATGTCATCGGAACGAGAGTTTAATCAGT	428
Qy	112	---GlyAsnAspLeuLeuGlnArgLeuLeuLeuValAspAspValGluArgLeuGlyIle	130
Db	429	CCCCACAATGATCTCTCAATCAACGCGTTTGGATGGTGTGATAGTGTGAACCGTTTAGGGATA	488



Query Match: 75.45% Indels: 16  
 DB: 6 Gaps: 6  
 US-10-025-145A-65 (1-618) x BD227661 (1-2018)

QY 1 MetAlaLeuSerIleThrProLeuValSerArgSerCysLeu----- 15  
 DB 6 ATGGCTCTAGTTTCTACCGCAGCGTGGCTTCCAAATCATGCTGCACAAATCGTTGATC 65  
 QY 16 SerSerSerHisGluLysLeuAlaLeuArgArgThrIleProThrLeuGlyLysCysArg 35  
 DB 66 AGTCTCTACCATGAGCTTAAGGCTCTCTCTAGAACAAATCCAGCTCTAGGAATGAGTAGG 125  
 QY 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55  
 DB 126 CAGGGAATATCATCTCTCCATCAGCATGAGCTTACCACCGCTGTATCCCGATGAT 185  
 QY 56 SerValGluArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIleGln 75  
 DB 186 GGTGTACGAAGACGCATGGCGATTTCCATTCCAACCTCTGGGACGATGATCATACAG 245  
 QY 76 SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95  
 DB 246 TCTTTA---CCAACGCTTATGAGGAAAAATCGTACTGGAGCGTCTCAGAACTGATC 302  
 QY 96 GlyGluValLysAspIleMetPheAsnPheLysSerLeuLysAspGlyGly----- 112  
 DB 303 GGGGAAGTAAAGAAC---ATGTTCAATTCGATGTCATTAGAGATGGAGATTGATGAT 359  
 QY 113 -----AsnAspLeuLeuGlnArgLeuLeuValAspValGluArgLeuGlyIle 130  
 DB 360 CCGCTCAATGATCTCATCAACGCTTTGGATTGTGACAGCGCTTGAAACGTTTGGGATC 419  
 QY 131 AspArgHisPheLysLysGluLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150  
 DB 420 CATAGACATTTCAAGATGAGATAAAATCGGCGCTTGATTATGTTTACAGTTATTGGGGC 479  
 QY 151 GluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170  
 DB 480 GAAATGGCATCGATCGCGAGGGAGAGTGTGTACTGAUTCAGTCAACTCACTGCGTTG 539  
 QY 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190  
 DB 540 GGCTTCGAACCTCAGCTACACGATACCGGTGCTTCAGATGTTTCAAAGCTTTC 599  
 QY 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyIleArg 210  
 DB 600 AAAGGCCAAATGGGCGAGTTTCCCTGCTCTGAAATATTTCAGACAGATGAAGAGATCAGA 659  
 QY 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp 230  
 DB 660 GCGTTCGAATTTATTCGGGCTCCCTCATTCGCTTCCAGGGGAGAAATATGATGAT 719  
 QY 231 GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer 250  
 DB 720 GAGGCTGAAATCTCTCTACCAATATTTAAAGAGAGCCCTGCAAAAGATTCGCGTCTCC 779  
 QY 251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTrpHisThrAsnLeuPro 270  
 DB 780 AGT---CTTTCGCGAGAGATCGGGGACGTTTGGATATGTTGGCACACATATTTGGCG 836  
 QY 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsnLysAsnAla 290  
 DB 837 CGATTGGAGCAAGGAATATACATCCAAAGTCTTTGGACAGGACACTGAGAACAGAGTCA 896  
 QY 291 -----AlaGluLysLeuLeuGluLeuAlaLysLeuGluPheAsnIlePheHisSer 307  
 DB 897 TATGTGAAGACGCAAAACCTTTTAGAACTCGCAAAATTTGAGTTCAACATCTTTCAATCC 956  
 QY 308 LeuGlnGluArgGluLeuLysHisValSerArgTyrTrpLysAspSerGlySerProGlu 327  
 DB 957 TTACAAAGAGGGAGTTAGAAAGTCTGTCAGATGTTGGAAAGAAATCGGCTTCTCTGAG 1016

QY 328 MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347  
 DB 1017 ATGACCTTCTCGGACATCGTCACTGGAAATACATACACTTGGCTTCTGCTGATCGGTTTC 1076  
 QY 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisIleuIleThrVal 367  
 DB 1077 GAGCCTCAACATCTCTGGATTCAAGCTCGGCTTGGCAAGACGCTCATCTTATCAAGGTT 1136  
 QY 368 LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387  
 DB 1137 CTTCAGCATATATGACACACCTTCGACAGTAGAGCTGGAACCTTCTTCAGGACACA 1196  
 QY 388 IleLysArgTyrAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407  
 DB 1197 ATGAAGAGATGGATCCGCTCGATAGATTCCTTCCAGATATATATGAAGAGAGTGTAC 1256  
 QY 408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysIleGlnGlyArg 427  
 DB 1257 ATAGCGGTTTACGACACCGTAAATGAAATGGCTTCGAGAGCGAGAGGCTCAAGGCCGA 1316  
 QY 428 AspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGlu 447  
 DB 1317 GATAGCTCACATATGCTCGGAGCTTCGGAGGCTTATATGATTCGTATATGCAAGAA 1376  
 QY 448 AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys 467  
 DB 1377 GCAAGGTGATCGCCACTGGTTTACCTGCCCTCCTTTGATGAGTACTACGAGAATGGAAA 1436  
 QY 468 ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe 487  
 DB 1437 GTTAGCTGTGGTCATCGCATATCCGCATTCGCAACCAATCTGCATATGACATCCCTTT 1496  
 QY 488 ProAspHisIleLeuLysGluValAspPheProSerLeuAsnAspLeuIleCysIle 507  
 DB 1497 CTGATCATATCTCAAGGAAGTTGACTTCCCATCAAAAGCTTACAGCTTGCATGTGCC 1556  
 QY 508 IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu 527  
 DB 1557 ATCCTTCGATTACGAGGTGATACGCGTCTCAAGCGGACAGAGGCTCGTGAGAAGAA 1616  
 QY 528 AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu 547  
 DB 1617 GCTTCTCTATATCATGTTATATGAAAGACATCTCTGGAGTATCAGAGGAAGATGCTCTC 1676  
 QY 548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLys 567  
 DB 1677 GATCATATCAAGCCCATGATCAGTACGTAATCAAGGATTAATTTGGAACTTCTCMAA 1736  
 QY 568 ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTrp 587  
 DB 1737 CCAGACATCAATGTTCCCATCTCGGCGAAGAAACATGCTTTTGACATCGCCAGAGCTTC 1796  
 QY 588 HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu 607  
 DB 1797 CATTACGGCTACAAATACCGACGCTACAGCGTTGCCAACCGTTGCAACGAGAGAGTTTG 1856  
 QY 608 ValMetArgThrValIleGluProValProLeu 618  
 DB 1857 GTCACGAGAACCTCTCTGAACTGTGCTTTC 1889

RESULT 5  
 AR222097  
 LOCUS  
 DEFINITION Sequence 3 from patent US 6429014.  
 ACCESSION AR222097  
 VERSION AR222097.1 GI:23329471  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Unclassified.  
 Unclasseified.  
 1 (bases 1 to 2018)  
 Steele,C.L., Bohlmann, J. and Croteau,R.B.  
 Monoterpene synthases from grand fir (Abies grandis)

JOURNAL Patent: US 6429014-A 3 06-AUG-2002;  
 FEATURES Location/Qualifiers  
 source 1. .2018  
 /organism="unknown"  
 /mol\_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 2,34e-190 Length: 2018  
 Score: 2453.00 Matches: 475  
 Percent Similarity: 84.63% Conservative: 59  
 Best Local Similarity: 75.28% Mismatches: 81  
 Query Match: 75.45% Indels: 16  
 DB: 6 Gaps: 6

US-10-025-145a-65 (1-618) x AR222097 (1-2018)

Qy 1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15  
 Db 6 ATGGCTCTAGTTCTACCGCACCGTTGGCTTCCAAATCATCGCTGCACAAATCGTTGATC 65  
 Qy 16 SerSerSerHisGluIleAlaLeuArgThrIleProThrLeuGlyIleCysArg 35  
 Db 66 AGTTCTACCATGAGCTTAAGCTCTCTCTAGAACAAATTCAGCTCTAGGAATGATAGG 125  
 Qy 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55  
 Db 126 CGAGGGAATATATCATCTCTCCATCAGCATGAGCTCTACACCGTTGTAAACGATGAT 185  
 Qy 56 SerValGlnArgArgValGlyAsnTyHisSerAsnLeuTrpAspAspPheIleGln 75  
 Db 186 GGTGTACGAACCATGGGCGATTTCCATTCCAACCTCTGGGACGATGATGTATACAG 245  
 Qy 76 SerLeuIleSerThrProTyGlyAlaProAspTyArgGluArgAlaAspArgIleIle 95  
 Db 246 TCTTTA--CCAACGGCTTAAGAGAAAATCGTACCTGGAGCGTGTGAGAACTGATC 302  
 Qy 96 GlyGluValLysAspIleMetPheAsnPhelYssSerLeuGluAspGlyGly----- 112  
 Db 303 GGGGAAGTAAAGAAC--ATGTTCAATTCGATGTCTATTAGAAAGATGGAGAGTTAATGAGT 359  
 Qy 113 -----AsnAspLeuLeuArgLeuLeuValAspValGluArgLeuGlyIle 130  
 Db 360 CCGCTCAATGATCATTCACGGCTTTGGAATGTGCACACGCTTGAACTGGGGATC 419  
 Qy 131 AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyArgValAsnSerTyTrpAsn 150  
 Db 420 CATAGACATTTCAAAGATGAGATAAATCGCGCTTGATTAATGTTTACAGTTATTGGGC 479  
 Qy 151 GluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170  
 Db 480 GAAATCGCATCGGATCGGGAGGAGAGTGTGTACTGATCTGAATCAACTCAACTGCGTTG 539  
 Qy 171 GlyLeuArgThrLeuArgLeuHisGlyTyThrValSerSerAspValLeuAsnValPhe 190  
 Db 540 GGGCTTGAACCTACGACTACAGGATACCCGTGCTTCAGATGTTTCAAAGCTTTC 599  
 Qy 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGlyGlyIleArg 210  
 Db 600 AAAGGCCAAAATGGCAGTTTTCTGCTCTCAAAATATTACAGACAGATGAAGATCAGA 659  
 Qy 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp 230  
 Db 660 GCGCTTCTGAATTTATTCCGGGCGCTCCCTCATTCGCTTTCCAGGGGAGAAAATATGATG 719  
 Qy 231 GluAlaGluThrPheSerThrLysTyLeuArgGluAlaLeuGlnLysIleProAlaSer 250  
 Db 720 GAGGCTGAAATCTTCTCTACCAAAATATTTAAAGAGCCCTGCAAAAGATTCGCGTCTCC 779  
 Qy 251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyArgIleThrHisThrAsnLeuPro 270  
 Db 780 AGT--CTTTCGCGAGAGATCGGGACGTTTTTGGAAATATGTTGGCACACATATTTCGGC 836

Qy 271 ArgLeuGluAlaArgAsnTyMetAspValPheGlyGlnHisThrLysAsnLysAsnAla 290  
 Db 837 CGATTGGAAGCAGAGGAATTACATCCAAGTCTTTGGACAGGACACTGAGAACACCAAGTCA 896  
 Qy 291 -----AlaGluLysLeuLeuGluAlaLysLeuGluPheAsnIlePheHisSer 307  
 Db 897 TATGTGAAGAGCAAAAACATTTTAGAACTCGCAAAATTTGGAGTTCAACATCTTTCAATCC 956  
 Qy 308 LeuGlnGluArgGluLeuLysHisValSerArgTrpLysAspSerGlySerProGlu 327  
 Db 957 TTACAAAGAGGGAGTTAGAAAGTCTGGTCAGATGGTGGAAAGAAATCGGTTTCTCTGAG 1016  
 Qy 328 MetThrPheCysArgHisArgHisValGluTyTrpAlaLeuAlaSerCysIleAlaPhe 347  
 Db 1017 ATGACCTTCTGCCGACATCGTCACGTGGAATACTACATTTGGCTTCTGCAATTCGGTTC 1076  
 Qy 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal 367  
 Db 1077 GAGCTTCAACATCTGGATTGAGACTCGGCTTTGCCAAGACGTGTCACTTATCACGGTT 1136  
 Qy 368 LeuAspAspMetTyArgValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387  
 Db 1137 CTTGACGATATGTACGACACCTTCGGCACAGTAGACGAGCTGGAACCTTCCACAGGACA 1196  
 Qy 388 IleLysArgTrpAspProSerAlaMetGluCysLeuProGluTyTrpMetLysGlyValTy 407  
 Db 1197 ATGAAGAGATGGGATCGTCTCGATAGATTGCTTCAGAAATATATGAAAGAGAGTGATC 1256  
 Qy 408 MetMetValTyHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427  
 Db 1257 ATACGGTTTACGACACCGTAATGAATGGCTCGAGAGGAGAGGCTCAAGGCCGA 1316  
 Qy 428 AspThrLeuAsnTyAlaArgGlnAlaTrpGluAlaCysPheAspSerTyMetGlnGlu 447  
 Db 1317 GATACGCTCATATGCTCGGGAAGCTTGGAGGCTTATATTGATTCGTATATCAAGAA 1376  
 Qy 448 AlalysTrpIleAlaThrGlyTyLeuProThrPheGluGluTyTrpLeuGluAsnGlyLys 467  
 Db 1377 GCAAGTGGATCGCACTGGTTACTGCTCTCTTGTATGAGTACTACGAGATGGGAAA 1436  
 Qy 468 ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProphe 487  
 Db 1437 GTTAGCTGTGGTCAATCGCATATCCGCAATTCGCAATTCGCAATTCGCAATTCGCTTT 1496  
 Qy 488 ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle 507  
 Db 1497 CCTCATATCTCTCAAGGAAGTTGACTTCCCATCAAAGCTTAACGACTTGGCATGTGCC 1556  
 Qy 508 IleLeuArgLeuArgGlyAspThrArgCysTyTrpLysAlaAspArgAlaArgGlyGlu 527  
 Db 1557 ATCTTCGATTACAGGATGATACGCGTGTCTACAAAGCGGACAGGCTCGTGGAGAGAA 1616  
 Qy 528 AlaSerSerIleSerCysTyMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu 547  
 Db 1617 GCCTCTCTATATCATGTTATATGAAGAACAATCTCGAGTATCAGAGGAGATGCTCTC 1676  
 Qy 548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLys 567  
 Db 1677 GATCATATCAACGCCATGATCAGTGACGTAATCAAAGGATTAATTTGGAACTTCTCAAA 1736  
 Qy 568 ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTrp 587  
 Db 1737 CCAGACATCAATGTTTCCCATCTCGCGGAAGAAACATGCTTTTGCATCGCAGAGCTTTC 1796  
 Qy 588 HisHisGlyTyArgTyArgAspGlyTySerPheAlaAsnValGluThrLysSerLeu 607  
 Db 1797 CATTAAGGCTTACAAATACCGAGACGCTACAGCGTTCGCAACGTTGAAACGAGAGTTG 1856  
 Qy 608 ValMetArgThrValIleGluProValProLeu 618  
 Db 1857 GTCACGAGAACCTCTCTTGAATCTGTGCTTGTG 1889

RESULT 6

AR240696  
LOCUS AR240696 2018 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 19 from patent US 6468772.  
ACCESSION AR240696  
VERSION AR240696.1 GI:27285845  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 2018)  
AUTHORS Chappell, J., Manna, K.R., Noel, J.P. and Starks, C.M.  
TITLE Methods of making modified polypeptides  
JOURNAL Patent: US 6468772-A 19 22-OCT-2002;  
FEATURES  
Location/Qualifiers  
1..2018  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores: 2,34e-190 Length: 2018  
Pred. No.: 2453.00 Matches: 475  
Score: 84.63% Conservative: 59  
Percent Similarity: 75.28% Mismatches: 81  
Best Local Similarity: 75.45% Indels: 16  
Query Match: 6 Gaps: 6  
DB: 6

US-10-025-145A-65 (1-618) x AR240696 (1-2018)

QY 1 MetAlaLeuSerIleThrProLeuValSerArgSerCysLeu----- 15  
Db 6 ATGGCTCTAGTTTCTACCGCAGCGTTGGCTTCCAAATCATGCTGCACAAATCGTTGATC 65  
QY 16 SerSerSerHisGluIleLysAlaLeuArgThrIleProThrLeuGlyIleCysArg 35  
Db 66 AGTTCTACCCATGAGCTTAAGGCTCTCTCTAGAACAAATCCAGCTCTAGAAATGAGTAGG 125  
QY 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55  
Db 126 CGAGGGAATCTATCATCTCTCCATCAGCATGAGCTTACCACCGCTGTGAACCGATGAT 185  
QY 56 SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIleGln 75  
Db 186 GGTGTACGAAGACGATGCGGCGATTTCCATTCCTCAACCTCTGGACGATGATGTCATACAG 245  
QY 76 SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95  
Db 246 TCTTTA---CCACGCGCTTATGAGGAAATCGTACTCGGAGCGTCTGAGAACTGATC 302  
QY 96 GlyGluValLysAspIleMetPheAsnPheLysSerLeuGluAspGlyGly----- 112  
Db 303 GGGGAAGTAAAGAAC--ATGTTCAATTCGATGTCATTAGAAGATGGAGAGTTAATGAGT 359  
QY 113 -----AsnAspLeuGlnArgLeuLeuValAspValAspValGluArgLeuGlyIle 130  
Db 360 CCGCTCAATGATCTCATTCACGCGCTTTGGATGTGCGACGCGCTTGAACGTTTGGGGATC 419  
QY 131 AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150  
Db 420 CATAGACATTTCAAAGATGAGATAAATCGGCGCTTGATTAAGTTTACAGATTATGGGCG 479  
QY 151 GluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170  
Db 480 GAAATGGCATCGATCGCGAGGAGAGTGTGTTTACTGATCTGAACCTCAACTGCGGTTG 539  
QY 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190  
Db 540 GGGCTTCGAACCTTACGACTACCGGATCCCGGTGTCTTCAGATGTTTCAAAGCTTTC 599  
QY 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGlyGluIleArg 210  
Db 600 AAAGGCCAAATCGGCAGATTTCCTGCTCGAATAATATTCAGACAGATGAAGAGATCAGA 659

QY 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp 230  
Db 660 GGGGTTCTCAATTTATTCGGGCGCTCCCTCATTCGCTTCCAGGGGAGAAAATATTAGGAT 719  
QY 231 GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer 250  
Db 720 GAGGCTGAAATCTCTCTACCAAAATATTTAAAGAGAGCCCTGCAAAAGATTCCGGTCTCC 779  
QY 251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTyrHisThrAsnLeuPro 270  
Db 780 AGT---CTTTCGGGAGAGATCGGGACGTTTGGAAATATGGTTGGCACACATATTGCGG 836  
QY 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsnLysAsnAla 290  
Db 837 CGATTGGAAGCAAGGAATTTACATCCAAAGTCTTTGGACAGGACACTGAGAACACGAAGTCA 896  
QY 291 -----AlaGluLysLeuLeuGluAlaLysLeuGluPheAsnIlePheHisSer 307  
Db 897 TATGTGAAGAGCAAAAACCTTTTAGAACTCGCAAAATTTGGAGTTCAACATCTTCAATCC 956  
QY 308 LeuGlnGluArgGluLeuLysHisValSerArgTyrTrpLysAspSerGlySerProGlu 327  
Db 957 TTACAAAAGAGGAGTTAGAAAGTCTGGTCAGATGGTGAAGAATCGGGTTTCTCTGAG 1016  
QY 328 MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347  
Db 1017 ATGACCTTCTCGCGACATCGTCAGGTGAATACTACACTTTGGCTTCTCGATTGGCTTC 1076  
QY 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisIleThrVal 367  
Db 1077 GAGCCTCAACATTTCTGGATTACACTCGGCTTTGGCAAGACGCTGTCATCTTATCAGGTT 1136  
QY 368 LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387  
Db 1137 CTTGACGATATGACACACCTTTCGGCACAGTAGACGAGCTGGAACCTCTTCACAGGACA 1196  
QY 388 IleLysArgTyrAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407  
Db 1197 ATGAAGAGATGGATCCGCTCTCGATAGATTGCTTCCAGAATATATGAAGAGAGTGTAC 1256  
QY 408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427  
Db 1257 ATAGCGGTTTACGACACCGTAATGAATGGCTCGAGAGCGACAGAGGAGCTCAAGGCCGA 1316  
QY 428 AspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGlu 447  
Db 1317 GATACGCTCACATATGCTCGGGAAGCTTCGGAGGCTTATATTGATTCTGATATGCAAGAA 1376  
QY 448 AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys 467  
Db 1377 GCAAGGTGGATCGCCACTGGTTACCTGCGCTCTCTTTGATGAGTACTACAGAGAATGGAAA 1436  
QY 468 ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe 487  
Db 1437 GTTAGCTGTGGTCATCGCATATCCGCATTCGCAACCCATCTTCACAATGGACATCCCTTT 1496  
QY 488 ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle 507  
Db 1497 CCTGATCATATCTCAAGGAAGTTGACTTCCCATCAAGCTTAACGACTTGGCATGTGCC 1556  
QY 508 IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu 527  
Db 1557 ATCTCTTCGATTACGAGGTGATACGGGTGCTCAAGCGCGACAGGGCTCGTGAGAGAA 1616  
QY 528 AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu 547  
Db 1617 GCTTCTCTATATCATGTTATATGAAGACATCTCGGAGTATCAGAGGAAGATGCTCTC 1676  
QY 548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLys 567  
Db 1677 GATCATATCAACGCCCATGATGATGACGTAATCAAAGGATTAATTTGGGAATCTCTCAAA 1736  
QY 568 ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTrp 587

Db 1737 CCAGACATCAATGTTCCCATCTCGCGAAGAAACATGCTTTTGACATCGCCAGAGCTTTC 1796  
Qy 588 HishisGlyTyrArgTyrArgGlyTyrSerPheAlaAsnValGluThrLysSerLeu 607  
Db 1797 CATTACGGCTACAAATACCGAGACGGCTACAGCGTTGCCAACGTTGAAACGAGAGTTTG 1856

Qy 608 ValMetArgThrValIleGluProValProLeu 618

Db 1857 GTCACGAGAACCTTCCTGATCTGTCCTTTG 1889

## RESULT 7

LOCUS AR266971 2018 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 19 from patent US 6495354.  
ACCESSION AR266971

VERSION AR266971.1 GI:29696426

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 2018)

AUTHORS Chappell,J., Manna,K.R., Noel,J.P. and Starks,C.M.

TITLE Synthesizes

JOURNAL Patent: US 6495354-A 19 17-DEC-2002;

FEATURES Location/Qualifiers

source 1. 2018

/organism="unknown"

/mol\_type="genomic DNA"

## ORIGIN

## Alignment Scores:

Pred. No.: 2,346-190 Length: 2018  
Score: 2453.00 Matches: 475  
Percent Similarity: 84.63% Conservative: 59  
Best Local Similarity: 75.28% Mismatches: 81  
Query Match: 75.45% Indels: 16  
DB: 6 Gaps: 6

US-10-025-145A-65 (1-618) x AR266971 (1-2018)

Qy 1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15  
Db 6 ATGGCTCTAGTTTACCGCACCGTGGCTTCCAAATCATGCTGCACAAATCGTTGATC 65  
Qy 16 SerSerHisGluIleLysAlaLeuArgArgThrIleProThrLeuGlyIleCysArg 35  
Db 66 AGTCTACCCATGAGCTTAAGGCTCTCTAGAACAAATCCAGCTTAGGAATGAGTAGG 125  
Qy 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55  
Db 126 CGAGGAAATCTATCACTCTCTCCATCAGCATGAGCTCTACCACCGTTGTAAACCGAT 185  
Qy 56 SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuThrProAspAspPheIleGln 75  
Db 186 GGTGTACGAAGACGATGGCGATTTCCTTCCAACTCTGGACGCTGCTGAGAACTGATC 245  
Qy 76 SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95  
Db 246 TCTTTA---CCAACCGCTTATGAGGAAATCGTACCTGGACGCTGCTGAGAACTGATC 302  
Qy 96 GlyGluValLysAspIleMetPheAsnPhelysSerLeuGluAspGlyGly----- 112  
Db 303 GGGGAAGTAAGAAC---ATGTTCAATTCGATGTCTATTAGAAAGATGGAGATTAATGAT 359  
Qy 113 -----AsnAspLeuGlnArgLeuLeuValAspValGluArgLeuGlyIle 130  
Db 360 CGCTCAATGATCTCATTCACACGCTTTGGATTGCGACGCTTGAACGTTGGGGATC 419  
Qy 131 AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150  
Db 420 CATAGACATTTCAAGATGAGATATAATCGCGCTTGATTATGTTTACAGTTATTGGGCG 479

Qy 151 GluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170  
Db 480 GAAATGGCATCGATCGGAGGAGGAGTGTCTTACTGATCTGAATCAACTCACTCGGTG 539  
Qy 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190  
Db 540 GGGCTTCGAACCCCTACGACTACCGGATACCCGGTGTCTTCAGATGTTTCCAAAGCTTTC 599  
Qy 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGluIleArg 210  
Db 600 AAGGCCCAAAATGGCAGTTTCTCTGCTCTGAAATATTCAGACAGATGAAGATCAGA 659  
Qy 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp 230  
Db 660 GGCCTTCTGAATTTATTCGCGGCTCCCTCATTCCTTCCAGGGGAGAAAATATGATGAT 719  
Qy 231 GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer 250  
Db 720 GAGCTGAAATCTTCTCTACCAATATTTAAAGAACCCCTGCAAAAGATTCCGGTCTCC 779  
Qy 251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTyrHisThrAsnLeuPro 270  
Db 780 AGT---CTTTCGCGAGAGATCGGGAGCTTTTGGAAATATGTTGGCACACATATTTCGCG 836  
Qy 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsnLysAsnAla 290  
Db 837 CGATTGGAAGCAAGGAATTACATCCAAGTCTTTGACAGGACACTGAGAACACGGAAGTCA 896  
Qy 291 -----AlaGluLysLeuLeuGluLeuAlaLysLeuGluPheAsnIlePheHisSer 307  
Db 897 TATGTGAAGACAAAATACTTTTAGACTCGCAAAATTTGGAGTTCAACATCTTTCAATCC 956  
Qy 308 LeuGlnGluArgGluLeuLysHisValSerArgTrpIlePheLysAspSerGlySerProGlu 327  
Db 957 TTCAAAAAGAGGGAGTTAGAAAGTCTGGTCAGATGGTGGAAAGAAATCGGGTTTCTCGAG 1016  
Qy 328 MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347  
Db 1017 ATGACCTTCTCGCGACATCGTCAGTGGAAATCTACACTTTGGCTTCTCGCATTCGCTTC 1076  
Qy 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal 367  
Db 1077 GAGCTCTCAACATTCGGAATTCAGACTCGGCTTGGCAAGACGTCATCTTATCAGGTT 1136  
Qy 368 LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387  
Db 1137 CTTGACGATATGTACGACACCTTCGGCAGTAGACGAGCTGGAACTCTTCACACGGACA 1196  
Qy 388 IleLysArgTrpAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407  
Db 1197 ATGAAGAGATGGGATCGCTCGCATAGATTGGCTTCCAGAAATATATGAAGGAGTGATC 1256  
Qy 408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427  
Db 1257 ATAGCGGTTTACGACACCGTAAATGAAATGGCTCGAGAGGCGAGGAGGCTCAAGCCGA 1316  
Qy 428 AspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGlu 447  
Db 1317 GATACGCTCACATATGCTCGGAGAGCTTGGGAGGCTTATATTGATTCGTATATGCAAGAA 1376  
Qy 448 AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys 467  
Db 1377 GCAAGTGGATCGGCACCTGGTTACTCGCCCTCTTGTAGTAGTACTACGAGAATGGGAAA 1436  
Qy 468 ValSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe 487  
Db 1437 GTAGCTGTGTCATCGCATATCCGATTCGCAATTCGACATTCGACATCGCATCCCTTT 1496  
Qy 488 ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle 507  
Db 1497 CTTGATCATATCTCAAGGAAGTTGACTTCCCATCAAGCTTAACGACTTGCATGTGCGC 1556  
Qy 508 IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu 527

```

1557 ATCTTCGATTACGAGTGATACGCGTGCTACAAGCGCGACAGGGCTCGTGAGAGAA 1616
Db
528 AlasSerSeriLeSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu 547
Qy
1617 GCCTTCCTCTATATCATGTTTATATGAAGAACAATCCCTGGAGTATCAGAGGAAGATGCTCTC 1676
Db
548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLys 567
Qy
1677 GATCATATACACCCATGATCAGTACGTAAATCAAGGATTAATTTGGGAACCTTCTCAA 1736
Db
568 ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTrp 587
Qy
1737 CCAGACATCAATGTTCCCATCTCGCGAAGAAACATGCTTTTGACATCGCCAGAGCTTTC 1796
Db
588 HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu 607
Qy
1797 CATTACGGCTACAAATACCGAGACGGCTACAGCGTTGCCAACGTTGAAACGAAGAGTTTG 1856
Db
608 ValMetArgThrValIleGluProValProLeu 618
Qy
1857 GTCAAGAGAACCCTCCTTGAATCTGTGCTTTG 1889
Db

RESULT 8
AR316320
LOCUS
DEFINITION
ACCESSION AR316320
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 (bases 1 to 2018)
AUTHORS Chappell,J., Manna,K.R., Noel,J.P. and Starks,C.M.
TITLE Synthases
JOURNAL Patent: US 6559297-A 19 06-MAY-2003;
FEATURES
Location/Qualifiers
1..2018
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 2,34e-190 Length: 2018
Score: 2453.00 Matches: 475
Percent Similarity: 84.63% Conservative: 59
Best Local Similarity: 75.28% Mismatches: 81
Query Match: 75.45% Indels: 16
DB: 6 Gaps: 6

US-10-025-145A-65 (1-618) x AR316320 (1-2018)

Qy 1 MetAlaLeuSerIleThrProLeuValSerArgSerCysLeu----- 15
Db 6 ATGGCTCTAGTTTCTACCCACCGCTTGGCTTCCAAATCATGCTCGACAAATCGTTGATC 65
Qy 16 SerSerSerHisGluIleLysAlaLeuArgArgThrIleProThrLeuGlyIleCysArg 35
Db 66 AGTTCTACCATCAGCTTAAGGCTCTCTAGACAAATTCAGCTTAGAATGAGTAGG 125
Qy 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55
Db 126 CCAGGGAATCTATCTACTCTCCATCAGCATGAGCTCTACCACCGCTTGAACCGATGAT 185
Qy 56 SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIleGln 75
Db 186 GGTGTACGAGACCGATGGCGATTTCCATTCACACCTCTGGACCATGATGTATACAG 245
Qy 76 SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95
Db 246 TCTTTA---CCAACGGCTTATGAGGAAAAATCGTACCTGGAGCGTGCTGAGAAACTGATC 302

```

```

Qy 96 GlyGluValLysAspIleMetPheAsnPheLysSerLeuGluAspGlyGly----- 112
Db 303 GGGGAGTAGAAGAAC---ATGTTTCAATTCGATGTCATTAGAAGATGAGAGTTAATGAGT 359
Qy 113 -----AsnAspLeuLeuGlnArgLeuLeuValAspAspValGluArgLeuGlyIle 130
Db 360 CCGCTCAATGATCTCATTCACCGCTTGGATTGTGCACAGCCTTGAACGTTGGGGATC 419
Qy 131 AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150
Db 420 CATAGACATTTCAAAGATGAGATAAAATCGCGCTTGATTTATGTTTACAGTTATTGGGC 479
Qy 151 GluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170
Db 480 GAAATGGCATCGATCGGGAGGAGAGTGTGTACTGATCTGAACCTCAACTGCGGTG 539
Qy 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190
Db 540 GGGCTTCGAACCTCAGACTACAGCATACCGGTGTCTTCAGATGTTTCAAAGCTTTC 599
Qy 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGluIleArg 210
Db 600 AAAGGCCAAATGGGAGTCTTCTGCTCTGAAAATATTTCAGACAGATGAAGAGATCAGA 659
Qy 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp 230
Db 660 GCGGTTCTGAATTTATTCGGGCTCTCCCTCATTCCTTTCCAGGGGAGAAAATATTGGAT 719
Qy 231 GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer 250
Db 720 GAGGCTGAAATCTTCTCTACCAAAATATTTAAAGAAAGCCCTGCAAAAGATTCGGGTCTCC 779
Qy 251 SerIleLeuSerLeuGluLeuArgAspValLeuGluTyrGlyTyrTrpHisThrAsnLeuPro 270
Db 780 AGT---CTTTCCGCGAGAGATCGGGAGCTTTTGGAAATATGTTGGACACATATTTGCCG 836
Qy 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsnLysAsnAla 290
Db 837 CGATTGGAAGCAAGGAATTACATCCAAAGTCTTTGGACAGGACACACTGAGAACAGAGTCA 896
Qy 291 -----AlaGluLysLeuLeuGluLeuAlaLysLeuGluPheAsnIlePheHisSer 307
Db 897 TATGTGAGAGCAAAAACTTTTAGAACTCGCAAAATTCGAGTTCAACATCTTCAATCC 956
Qy 308 LeuGlnGluArgGluLeuLysHisValSerArgTyrTrpTrpLysAspSerGlySerProGlu 327
Db 957 TTACAAAAGAGGAGGTAGAAAGTCTGGTCAGATGGTGGAAAGAAATCGGGTTCCTGAG 1016
Qy 328 MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347
Db 1017 ATGACCTTCTGCCGACATCGTCACGTGGAATACACATTTGGCTTCCTCGATGGGTTC 1076
Qy 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal 367
Db 1077 GAGCCTCAACATCTCTGGATTACACTCGGCTTTGGCAAGACGTGTCTATTCACGGTT 1136
Qy 368 LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387
Db 1137 CTTGACGATATGTACGACACCTTCGGCAGCATGAGAGCTGGAACCTCTTCACAGGACA 1196
Qy 388 IleLysArgTyrAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407
Db 1197 ATGAAGAGATGGATCGGCTCCGTAGATTCGCTTCCAGAAATATATGAAGAGGTGATC 1256
Qy 408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427
Db 1257 ATAGCGGTTTACGACACCGTAATGAAATGGCTTCGAGAGCAGAGAGGCTCAAGGCCGA 1316
Qy 428 AspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGlu 447
Db 1317 GATACGCTCACATATGCTCGGAGAGCTTGGGAGGCTTATATGATTCGTATATGCAAGAA 1376
Qy 448 AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys 467

```



```

Db 1197 ATCAAGAGATGGATCGCTCTCGATAGATTCCTCCAGATATATGAAGAGGTGTAC 1256
QY 408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427
Db 1257 ATAGCGGTTTACGACACCCGTAATGAAATGGCTTCGAGAGCGAGAGGCTCAAGGCCGA 1316
QY 428 AspThrLeuAsnTyrAlaArgGlnAlaTyrGluAlaCysPheAspSerTyrMetGlnGlu 447
Db 1317 GATACGCTCACATATGCTCGGAAGCTTGGAGGCTTATGATTCTGATATGCAAGAA 1376
QY 448 AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys 467
Db 1377 GCAAGGTGATCGCCACTGGTTACCTGCCCTCTTGTGATGAGTACTACGAGATGGAAA 1436
QY 468 ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe 487
Db 1437 GTTAGCTGTGGTTCATCGCATATCCGCAATTCGCAACCCATCTGACAAATGGACATCC 1496
QY 488 ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle 507
Db 1497 CTTGATCATATCTCAAGAAAGTTGACTTCCCATCNAAGCTTAAAGACTTGGCATGTGCC 1556
QY 508 IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu 527
Db 1557 ATCTTCGATTACGAGGTGATACGGGTGCTACAGCGCGGACAGGGCTCGTGGAGAAGAA 1616
QY 528 AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu 547
Db 1617 GCTTCTCTATATCATGTATATGTAAGAAACAATCTCGGAGTATCAGAGGAAGATGCTCTC 1676
QY 548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLeuLys 567
Db 1677 GATCATATCAAGCCATGATCATGACGTAATCAAGAGATTAAATTTGGAACTTCTCAA 1736
QY 568 ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTrp 587
Db 1737 CCAGACATCAATGTTCCCATCTCGCGGAAGAAACATGCTTTTGACATCCCGAGAGCTTC 1796
QY 588 HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu 607
Db 1797 CATTACGGGTACAAATACCGAGACGGCTACACGGTTCACACGTTGAAACGAAGAGTTTG 1856
QY 608 ValMetArgThrValIleGluProValProLeu 618
Db 1857 GTACGAGAACCTCTCTGAACTGTGCTTTG 1889

```

```

RESULT 10
AR429869
LOCUS AR429869 2018 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 19 from patent US 6645762.
ACCESSION AR429869
VERSION AR429869.1 GI:40190267
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

```

```

REFERENCE
1 (bases 1 to 1818)
AUTHORS Chappell,J., Manna,K.R., Noel,J.P. and Starcks,C.M.
TITLE Syntheses
JOURNAL Patent: US 6645762-A 19 11-NOV-2003;
FEATURES
1. -2018
Source /organism="unknown"
/mol_type="genomic DNA"

```

```

Alignment Scores:
Pred. No.: 2,34e-190 Length: 2018
Score: 2453.00 Matches: 475
Percent Similarity: 84.63% Conservative: 59
Best Local Similarity: 75.28% Mismatches: 81
Query Match: 75.45% Indels: 16

```

```

DB: 6 Gaps: 6
US-10-025-145A-65 (1-618) x AR429869 (1-2018)
QY 1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15
Db 6 ATGGCTCTAGTTTCTACCGCACCGTTGGCTTCCAAATCATGCTGCACAAATCGTTGATC 65
QY 16 SerSerSerHisGluIleLysAlaLeuArgArgThrIleProThrLeuLysLysCysArg 35
Db 66 AGTTCTACCATGAGCTTAAGCTCTCTAGAAACAAATCCAGCTCTAGGAATAGTAGG 125
QY 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55
Db 126 CGAGGGAATCTATCACTCTTCATCAGCATGAGCTCTACACCGTTGTAACCGATGAT 185
QY 56 SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIleGln 75
Db 186 GGTGTACGAAGACGCGATGGCGATTTCATTCCAACTCTGGACGATGATGTCATACAG 245
QY 76 SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95
Db 246 TCTTTA---CCAACCGCTTATGAGAAATCGTACCTGGAGCGCTGAGAACTGATC 302
QY 96 GlyGluValLysAspIleMetPheAsnPheLysSerLeuGluAspGlyGly----- 112
Db 303 GGGGAGTAAGAAC---ATGTTCAATTCGATGTCATTAGAGATGAGAGTTAATGAGT 359
QY 113 -----AsnAspLeuGlnArgLeuLeuValAspValGluArgLeuGlyIle 130
Db 360 CGCGCTCAATGATCTCATTCACGCTTTGGATTGTTCGACAGCTTGAACGTTTGGGATC 419
QY 131 AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150
Db 420 CATAGACATTTCAAAGATGAGATAAATCGGCGCTTGATTATGTTTACAGTTATTGGGCG 479
QY 151 GluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170
Db 480 GAAATGGCATCGGATCGGGAGGAGAGTGTGTTACTGATCTGAACCTCAACTCGGTTG 539
QY 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190
Db 540 GGGCTTCGAACCTCAGACTACCGATACCGCGTGTCTTCAGATGTTTTCAAAGCTTTC 599
QY 191 LysAspLysAsnGlyGlnPheSerThrAlaAsnIleGlyIleGluGlyGluIleArg 210
Db 600 AAAGGCCAAATGGGCGAGTTTCTCTGCTCTGAAAATATTTCAGACAGATGAAGATCAGA 659
QY 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp 230
Db 660 GCGCTTCTGAATTTATTCGGGCGCTCCCTCATTTGCTTTCAGGGGAGAAAATATGATGAT 719
QY 231 GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGluLysIleProAlaSer 250
Db 720 GAGGCTGAAATCTCTCTACCAAAATATTAAAGAAAGCCCTGCAAAAGATTCGGGTCTCC 779
QY 251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTrpHisThrAsnLeuPro 270
Db 780 AGT---CTTTCCGAGAGATCGGAGCGTTTGGAAATATGTTGGACACATATTTGCGG 836
QY 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsnLysAsnAla 290
Db 837 CGATTGGAAGCAAGGAATTTACATCAAGTCTTTGGACAGGACACTGAGAACAGAGTCA 896
QY 291 -----AlaGluLysLeuLeuAlaLysLeuGluPheAsnIlePheHisSer 307
Db 897 TATGTGAAGAGCAAAAACCTTTTAGAACTCGCAAAATTTGGAGTTCAACATCTTCAATCC 956
QY 308 LeuGlnGluArgGluLeuLysHisValSerArgTyrTrpLysAspSerGlySerProGlu 327
Db 957 TTACAAAAGAGGAGGTAGAAAGTCTGGTCAGATGGTGGAGAAATCGGGTTTCTCTGAG 1016
QY 328 MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347

```

```
Db 1017 ATGACCTTCCCGACATCGTCACGTGAATACTACATTGGCTTCTCGCATTCGCGTTC 1076
Qy 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal 367
Db 1077 GAGCCTCAACATTCGTGATTCCAGACTCGGCTTCCAGAGAGCGTGCATCTTATCACGGTT 1136
Qy 368 LeuAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387
Db 1137 CTTGACCATATGTACGACACCTTCGGACACGTAGACGAGCTTGAACCTTCCACAGCGACA 1196
Qy 388 IleLysArgTyrAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407
Db 1197 ATGAGAGATGGATCCGCTCTCATAGATTGCTTCCAGAAATATATAAAGGAGTGATC 1256
Qy 408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427
Db 1257 ATAGCGGTTTACGACACCGTAAATGAATGCTCGAGAGGCGAGGAGGCTCAAGGCCGA 1316
Qy 428 AspThrLeuAsnTyrAlaArgGlnAlaTyrGluAlaCysPheAspSerTyrMetGlnGlu 447
Db 1317 GATACGCTCATATATGCTCGGAAAGCTTGGGAGGCTTATATTGATTCGTATATCAAGAA 1376
Qy 448 AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys 467
Db 1377 GCAAGGTGATCGGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1436
Qy 468 ValSerSerAlaHisArgProCysAlaLeuGlnProLeuLeuThrLeuAspIleProPhe 487
Db 1437 GTTAGCTGTGCTCATCGCATATCGCATTCGCAATTCGCAATTCGCAATTCGCAATTC 1496
Qy 488 ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle 507
Db 1497 CCTGATCATATCCTCAAGAAAGTTGACTTCCCATCAAGAGCTTACGACTTGGCATGTC 1556
Qy 508 IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu 527
Db 1557 ATCCTTCGATTACAGGTGATACCGGTGCTACAGCGGCGACAGGCTCGTGGAGAGAA 1616
Qy 528 AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu 547
Db 1617 GCTTCCTCTATATCATGTTATGAAGACAAATCTGGAGTATCAGAGGAGATGCTCTC 1676
Qy 548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLys 567
Db 1677 GATCATATCAACGCCATGATCAGTCAAGTCAATCAAGGATTAATTCGGAATCTTCTCAA 1736
Qy 568 ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTyr 587
Db 1737 CCAGACATCAATGTTCCCATCTCGCGGAGAGAAACATGCTTTTGACATCGCCAGAGCTTC 1796
Qy 588 HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu 607
Db 1797 CATTCAGGCTACAAATACCGAGACGGCTACAGCGTTGCCAACGTTGAAACGAGAGTTTG 1856
Qy 608 ValMetArgThrValIleGluProValProLeu 618
Db 1857 GTCACGAGAACCCCTCTGTAATCTGTGCTTTG 1889

RESULT 11
AGU87909
LOCUS AGU87909 2018 bp mRNA linear PLN 24-SEP-1997
DEFINITION Abies grandis pinene synthase (AG3.18) mRNA, complete cds.
ACCESSION U87909
VERSION U87909.1 GI:2411482
KEYWORDS .
SOURCE Abies grandis
ORGANISM Abies grandis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
REFERENCE 1 (bases 1 to 2018)
AUTHORS Bohlmann,J., Steele,C.L. and Croteau,R.
TITLE Monoterpene synthases from grand fir (Abies grandis). cDNA
```

```
isolation, characterization, and functional expression of myrcene
synthase, (-)-(4S)-limonene synthase, and (-)-(1S,5S)-pinene
synthase
J. Biol. Chem. 272 (35), 21784-21792 (1997)
97413772
PUBMED 9268308
REFERENCE 2 (bases 1 to 2018)
AUTHORS Bohlmann,J., Steele,C.L. and Croteau,R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-1997) Institute of Biological Chemistry,
Washington State University, Clark Hall, Pullman, WA 99164-6340,
USA
FEATURES
source
1..2018
/organism="Abies grandis"
/mol_type="mRNA"
/db_xref="taxon:46611"
1..2018
/gene="AG3.18"
6..1892
/gene="AG3.18"
note="monoterpene synthase; terpene synthase; monoterpene
synthase; terpene cyclase"
/codon_start=1
/product="pinene synthase"
/protein_id="AAB71085.1"
/db_xref="GI:2411483"
/translation="MALVSTAPLASKSLKHSLSSTHELKALSRTPALQMRRGKS
ITPNSMSSTVTVDGVRMRMGDFHNLMDDDVIOQLPTAYEKSYLEAEKLIQVEV
KNWNSLEDGELMSPLNDLIQRLWLVDSLERLGIHRHFKDELKSLDYVSYWGEN
GIGGRESVTDLNSALGRLTLRHGYPVSSDVFKAQKQNGQFSCSENITDDEEIR
GVLNIFRASLIAFPGEKIMDEAIFSTKYLKEALQKIPVSLSEIGDVLGYGHTYL
PRLEARNYIQVGGDTENTKSYVKSKLLEKLEFNIFQSLQKRELESLEVRWVKESG
FPWMTFCRHRHVEYITLASCIAPEQHSGLRGLFAGTCHLITVLDDMVDTPGTVDLE
LFTATMKRWDPSIDCLPEYMKGVYJAVDTVMEMAREAEACGRDRLTVAREAEAY
IDSVMQEARMIATGYPSEFVEYGVKSGHRISALQPIITMDIPPDHILKVDFF
SKNDLACAILRLNGDTRCYKADRGEEASISCYMKDNPGVSEEDALDHINAMISD
VTKGLNELLKPDINVPISAKHAFDIARAFHYGYKYRVDGYSVANVETKSLVTRTLE
SVPL"
ORIGIN
Alignment Scores: 2.34e-190 Length: 2018
Pred. No.: 2453.00 Matches: 475
Score: 84.63% Conservative: 59
Percent Similarity: 81
Best Local Similarity: 75.28% Mismatches: 16
Query Match: 75.45% Indels: 6
DB: 8 Gaps: 6
US-10-025-145A-65 (1-618) x AGU87909 (1-2018)
Qy 1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15
Db 6 ATGGCTCTAGTTCTTACCGCACCGCTTGGCTTCCAAATCATGCTGCACAAATCGTTGATC 65
Qy 16 SerSerSerHisGluIleLysAlaLeuArgArgThrIleProThrLeuGlyIleCysArg 35
Db 66 AGTTCTACCATGAGCTTAAGGCTCTCTCTAGAACAAATTCAGCTCTAGGAATGAGTAGG 125
Qy 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55
Db 126 CGAGGGAATCTATCATCTCTCCATCAGCATGAGCTCTACCACTGTTAACCAGATGAT 185
Qy 56 SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTyrAspAspPheIleGln 75
Db 186 GGTGTACGAGACGATGGCGGATTCATTCACCTCTGGGACGATGATGTCATACAG 245
Qy 76 SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95
Db 246 TCITTTA---CCAAACGGCTTATAGGAAAAATCGTACCTGGAGCGTCTGAGAAACTGATC 302
Qy 96 GlyGluValLysAspIleMetPheAsnPheLysSerLeuGluAspGlyGly----- 112
```

Db 303 GGGGAAGTAAAGAAC---ATGTTCAATTGCGATGCTATTAGAGATGCGAGAGTTAATGAGT 359

Qy 113 -----AsnAspLeuLeuGlnArgLeuLeuValAspValGluArgLeuGlyIle 130

Db 360 CGGCTCAATGATCTCATCAACGCCCTTGGATGTCGACAGCCCTTGAACGTTTGGGGATC 419

Qy 131 AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150

Db 420 CATAGACATTTCAAGATGAGATAAATCGGCGCTTGATATGTTTACAGATTATGGGCG 479

Qy 151 GluLysGlyIleCysGlyArgGluSerValThrAspLeuAsnSerThrAlaLeu 170

Db 480 GAAATGGCATCGGATCGGGAGGAGAGTGTGTTACTGATCTGAACCTCAACTGCGGTG 539

Qy 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190

Db 540 GGGCTTCGAACCTCAGACTACCGGATACCGGCTCTTCAGATGTTTCAAAAGCTTTC 599

Qy 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyIleArg 210

Db 600 AAGGCCAAATGGGCAGTTTCTCTCTGAAATATTCAGACAGATGAAGNATCAGA 659

Qy 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp 230

Db 660 GCGTTCCTGAATTTATTCGGGCTCCCTCAUATGCTTCCAGGGGAGAAAATATGGAT 719

Qy 231 GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer 250

Db 720 GAGGCTGAAATCTCTCTACCAATATTTAAAGAGAGCCCTGCAAAAGATTCGGTCTCC 779

Qy 251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTrpHisThrAsnLeuPro 270

Db 780 AGT---CTTTCGGAGAGATCGGGAGCGTTTGGATATGTTGGCACATATTTGCCG 836

Qy 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsnLysAsnAla 290

Db 837 CGATTGGAAGCAAGGAATATACATCCAAAGTCTTTGGACAGGACACTGGAACACGAAGTCA 896

Qy 291 -----AlaGluLysLeuLeuGluLeuAlaLysLeuGluPheAsnIlePheHisSer 307

Db 897 TATGTGAAGAGCAAAATCTTTAGAACTCGCAAAATGGAGTTCAACATCTTCAATCC 956

Qy 308 LeuGlnGluArgGluLysHisValSerArgTrpTrpLysAspSerGlySerProGlu 327

Db 957 TTACAAAGAGGAGGTAGAAAGTCTGGTCAGATGCTGGAAGAAATCGGTTTTCCTGAG 1016

Qy 328 MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347

Db 1017 ATGACCTTCGCGACATCGTCACGTGGAATACTACACTTTGGCTTTCCTGCATTCGGTTC 1076

Qy 348 GluProGluHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal 367

Db 1077 GAGCCTCAACATCTCGATTACACTCGGCTTTCGCAAGACGTGTCATCTTATCAGGTT 1136

Qy 368 LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387

Db 1137 CTTGACGATATGACGACACCTTCGCGACAGTAGACGAGCTGGAATCTTTCAGCGACA 1196

Qy 388 IleLysArgTrpAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407

Db 1197 ATGAAGAGATGGGATCGGCTCGATAGATGCTTCCAGAAATATATGAAGAGGTGTAC 1256

Qy 408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427

Db 1257 ATAGCGGTTTACGACCGGTAAATGAAATGGCTTCGAGAGCGAGAGGCTCAAGGCCGA 1316

Qy 428 AspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGlu 447

Db 1317 GATACGCTCATATGCTCGGGAAGCTTGGAGGCTTATATGATTGTTATATGCAAGAA 1376

Qy 448 AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys 467

Db 1377 GCAAGGTGGATCGCCACTGGTTACTCGCTCCCTCTTGTATGATGAGTACTACGAGAATGGAAA 1436

Qy 468 ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe 487

Db 1437 GTTAGCTGTGTCATCGCATATCCGATTCGCAACCACTTCTGACATGAGCATCCCTTT 1496

Qy 488 ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle 507

Db 1497 CCGTATCATATCTCAAGGAAGTTGACITTCCTCAATCAAGCTTAAACGACTGGCATGGC 1556

Qy 508 IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu 527

Db 1557 ATTCCTTCGATTACGAGGTGATACGCGGTGCTCAAGGGCGGACAGGGCTCGTGAGAGAA 1616

Qy 528 AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu 547

Db 1617 GTTCTCTATATCATGTTATATGAAGACATCTCTGGAGTATCAGAGAGAGATGCTCTC 1676

Qy 548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLys 567

Db 1677 TATCATATCAAGCCCATGATCAGTGACGTAATCAAGGATTAAATTGGGAATCTTCTCAA 1736

Qy 568 ProAsnSerValProIleThrSerLysHisAlaPheAspIleSerArgValTrp 587

Db 1737 CCAGACATCAATGTTCCCATCTCGGCGAAGAAACATGCTTTTGCATCCGACAGCTTTC 1796

Qy 588 HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu 607

Db 1797 CATACGGCTACAAATACCGACGCTACAGGTTGCCAAGTTGAACGAGAGATTTC 1856

Qy 608 ValMetArgThrValIleGluProValProLeu 618

Db 1857 GTCAAGAGAACCTCTCTGAATCTGTGCTTTG 1889

## RESULT 12

AF543527 2082 bp mRNA linear PLN 10-MAR-2003  
LOCUS Pinus taeda (-) -alpha-pinenene synthase mRNA, complete cds.  
DEFINITION AF543527  
ACCESSION AF543527  
VERSION AF543527.1 GI:28894481  
KEYWORDS  
SOURCE Pinus taeda (loblolly pine)  
ORGANISM Pinus taeda  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
REFERENCE 1 (bases 1 to 2082)  
AUTHORS Phillips M.A., Wildung, M.R., Williams, D.C., Hyatt, D.C. and  
Croteau, R.  
TITLE cDNA isolation, functional expression, and characterization of  
(+)-alpha-pinenene synthase and (-)-alpha-pinenene synthase from  
loblolly pine (Pinus taeda): Stereococontrol in pinene biosynthesis  
Arch. Biochem. Biophys. 411 (2), 267-276 (2003)  
JOURNAL 22510022  
MEDLINE 12623076  
PUBMED  
REFERENCE 2 (bases 1 to 2082)  
AUTHORS Phillips, M.A., Wildung, M.R., Williams, D.C. and Croteau, R.B.  
TITLE Direct Submission  
JOURNAL Submitted (09-SEP-2002) Institute of Biological Chemistry,  
Washington State University, PO Box 646340, Pullman, WA 99164-6340,  
USA

## FEATURES

## source

1. .2082  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/db\_xref="taxon:3352"  
32..1921  
/note="monoterpene synthase"

## CDS

/codon\_start=1  
/product="(-)-alpha-pinenene synthase"  
/protein\_id="AA061225.1"  
/db\_xref="GI:28894482"  
translation="MSPVSVISPSDLCLPTSIDRSGRELIPHLITIPNVAMRROCK  
LMTASMSNLTAVSDAVIRPGRDPSNLDLLOSLSPSPGSPSPYRERAILIG  
EVKSNFSNSNEDGESITPLDDLIORLWVDSVERLIGDRHPKKEIKSALDHVRYWS

BKIGCGRESVYTDLNSALGRLTLRLJGYSADVLNHPKQSGQSFACITLKOTQDQ  
 RVLNLRVSLAFJPGKQWDRSFSKYLKEALQKLPVSSFGREIGDVLVETGWYH  
 LPLREARNYIDVPGQTEKSSYMKTEKLELALQFIFHALQKQRLVLEWYKWS  
 GSPQFTCRHRVVEYTLASCLAFQHSQFRLGAKACHITVLDMDYDTFGLDEL  
 ELFTSAIKWMDPSATECLPEYKMGVYMYVINTVNSMQSQAQKQRLTQYCHQWEE  
 YIDAMQAEAKWISGEVPTFEYENKRGVSSGHRYSLAQPLITLDTPFPEHLKVEDI  
 YIDAMQAEAKWISGEVPTFEYENKRGVSSGHRYSLAQPLITLDTPFPEHLKVEDI  
 PSQLNDALALRLRGQYQADRAKGEAESICYNMKMGPTGTEDADHLNLNAMIS  
 DVIKGNWELLKPNSSGVPISAKKHAPDISRAPHCGYKRDGYSVANITETKSLVKRTVI  
 DPTVL"

## ORIGIN

Alignment Scores:

Pred. No.:	3.45e-180	Length:	2082
Score:	2328.50	Matches:	456
Percent Similarity:	82.62%	Conservative:	67
Best Local Similarity:	72.04%	Mismatches:	91
Query Match:	71.62%	Indels:	19
DB:	8	Gaps:	8

US-10-025-145A-65 (1-618) x AF543527 (1-2082)

Qy	1	MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeuSerSerSer	----- 18
Db	32	ATGTCCTCTGTTCTGTCGTCCTCCGACTTGTGCTGCCACCATCGTTTCATC	91
Qy	19	-----HisGluIleLysAlaLeuArgArgThrIleProThrLeuGlyIleCys	34
Db	92	GACAGGTCGTGGTGAGCTTATACCTCTCCATATAACAATTCCAAATGTCGAATGCCG	151
Qy	35	ArgProGlyLysSerValAlaHis---SerIleAsnMetCysLeuThrSerValAlaSer	53
Db	152	AGGCAAGGGAATTAATGACACGCTCTCCATGACATGAATTTGAGGACCGCGTATCT	211
Qy	54	ThrAspSerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPhe	73
Db	212	GATGATGCTGTATAAGACGCGAGAGTGATTTCCATTCCACTCTGGGACCATGATTTG	271
Qy	74	IleGlnSerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArg	93
Db	272	ATACAGTCCCTT---TCCTCGCCTTATGGGAACCTCTTATCGGGAACGTCTGAGAGA	328
Qy	94	LeuIleGlyGluValLysAspIleMetPheAsnPheLysSerLeuGluAspGlyGly---	112
Db	329	CTGATGGGGAAGTAAAGAAT---AGCTTCAATTCAAATGTCAAAGCATGGCGAATCA	385
Qy	113	-----AsnAspLeuLeuGlnArgLeuLeuValAspAspValGluArgLeu	128
Db	386	ATCACTCCCTCCGATGATCTGATTCACGCTTTTGGATGGTGCACAGTGTGAAACGTTTG	445
Qy	129	GlyIleAspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyr	148
Db	446	GGAAATCGACACATTTTCAAAAAGAGATAAATCAGCGCTCGAUCATGTTTACAGGTAT	505
Qy	149	TrpAsnGluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThr	168
Db	506	TGGAGCGAAAAGGCAATGGATGTGGGAGAGAGTGTGTTTACTGACCTCAACTCAACT	565
Qy	169	AlaIleGlyLeuArgThrLeuArgIleuHisGlyTyrThrValSerSerAspValLeuAsn	188
Db	566	GCCTTGGGTCTTCGAACCCCTTCGTGACGATACGATGTGTCGACACGTTTTGAAT	625
Qy	189	ValPheLysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGlu	208
Db	626	CATTTCAAAATCAAGTGGGAGATTGCTGCACCTCTGAAG---CAGACAGAGGATCAG	682
Qy	209	IleArgGlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysVal	228
Db	683	ATCAGAACTGTACTTAAATTTATATCGGCTTCCTCTCATTTCCAGGAGAGAAAGTC	742
Qy	229	MetAspGluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnIlyIlePro	248
Db	743	ATGGATGAGGCTGAAGATTTCTCTCGAAATATTTGAAAGAACCCCTGCAAAGATTCCA	802

Qy	249	AlaSerSerIleLeuSerLeuGluIleAlaArgAspValLeuGluTyrGlyTyrHisThrAsn	268
Db	803	GTCTCCAGT---TTTTTCAAGAGATAGGGGACGTCCTCGAGTATGGTTGGCACACGAT	859
Qy	269	LeuProArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsnLys	288
Db	860	TTGCACGATTGGACGACGAGAAATTACATCGACGTCTTCGACAGGACACTGAACAACAGC	919
Qy	289	AsnAla-----AlaGluLysLeuLeuGluAlaLysLeuGluPheAsnIlePhe	305
Db	920	AAGTCATATATGAAGACCGAAGAACTTCTCGAACTTGCCAAAGTTGGAGTTCAACATCTTT	979
Qy	306	HisSerLeuGlnGluArgGluLeuLysHisValSerArgTrpTrpLysAspSerGlySer	325
Db	980	CAGCCCTTACAAAGCGAGAGTTGGAAATATCTCGTTAGATCGTGGAAAGGCTCTGGTTCG	1039
Qy	326	ProGluMetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIle	345
Db	1040	CCTCAATGACCTTTTGTGCACATCTGTCACTGGAGTACTACACTTTGGCTTCTTGCATT	1099
Qy	346	AlaPheGluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIle	365
Db	1100	GGGTTTCAGCCTCAACATTCTGGATTTCAGACTCGGCTTGCCAAAGCCTGTCAATATCATC	1159
Qy	366	ThrValLeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThr	385
Db	1160	ACGGTCTTCATGATATGTACGACACCTTCGGAACACTTCAGCAGCTCGAACTCTTCACA	1219
Qy	386	AlaThrIleLysArgTrpAspProSerAlaMetGluCysLeuProGluTyrMetLysGly	405
Db	1220	TCGTCAATTAGAGATGGGATCCGTCGGCGACAGAGTGCCCTTCGAGAAATATATGAAAGGA	1279
Qy	406	ValTyrMetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGln	425
Db	1280	GTTTACATGATAGTTTACACACTGCTAAATGCAATGTCTCAGGAGGCGACCAAGGCTCAA	1339
Qy	426	GlyArgAspThrLeuAsnTyrAlaArgGluAlaTrpGluAlaCysPheAspSerTyrMet	445
Db	1340	GGCGGACAGCGCTCAACTATTGTGCAGAGCGCTTGGGAGGAATATATTGATCGCTATATG	1399
Qy	446	GluGluAlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsn	465
Db	1400	CAAGAACAAAGTGGATCGCGAGTGGTGAAGTGCCAACTTTGAGGAGTACTATGAGAAC	1459
Qy	466	GlyLysValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIle	485
Db	1460	GGGAAAGTTAGCTCTGCTCATCGGTATCGGCATTGCCAATTTTCAGCAGCCGACATC	1519
Qy	486	ProPheProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIle	505
Db	1520	CCCTTTCTCGAGCAGCGCTCAAGGAAGTTGACATTCATCGCAGCTCAATGACTTGGCA	1579
Qy	506	CysIleIleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGly	525
Db	1580	TCTGCCATCTTCGATTACGAGGGGATACGCGCTGTACCAGGCGGACAGGCGCCGTGGA	1639
Qy	526	GluGluAlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAsp	545
Db	1640	GAAGAAGCTTCGTGTATCTTGTATATGAAGACAACTCTTGAACAACACAGAGAGAGAT	1699
Qy	546	AlaLeuAsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeu	565
Db	1700	GCTCTCAATCATCTCAACGGCCATGATCAGTGTATTAATTAAGGATTAATAATTTGGAGCTT	1759
Qy	566	LeuLysProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArg	585
Db	1760	CTCAAAACCAACACAGCAGCGTTCCCATATCTGCCAAAAAACAATGCTTTTGACATTAGCAGA	1819
Qy	586	ValTrpHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLys	605
Db	1820	GCTTTCCATTGTGGCTCAAAATATCGAGATGGCTACAGCGTTGCCAACATTGAACAACAAAG	1879
Qy	606	SerLeuValMetArgThrValIleGluProValProLeu	618

```
Db 1880 AGTTTGGTGAAGAACCGTCATTGATCTGCTCACTTTA 1918
|||||
RESULT 13
AF543529 2162 bp mRNA linear PLN 10-MAR-2003
LOCUS Pinus taeda alpha-terpineol synthase mRNA, complete cds.
DEFINITION AF543529
ACCESSION AF543529
VERSION AF543529.1 GI:28894485
KEYWORDS Pinus taeda (loblolly pine)
SOURCE Pinus taeda
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
REFERENCE Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
AUTHORS Phillips,M.A., Wildung,M.R., Williams,D.C., Hyatt,D.C. and
Croteau,R.
TITLE cDNA isolation, functional expression, and characterization of
JOURNAL (+)-alpha-pinene synthase and (-)-alpha-pinene synthase from
MEDLINE loblolly pine (Pinus taeda): stereocontrol in pinene biosynthesis
PUBMED Arch. Biochem. Biophys. 411 (2), 267-276 (2003)
REFERENCE 22510022
2 (bases 1 to 2162)
2 (bases 1 to 2162)
2 (bases 1 to 2162)
AUTHORS Phillips,M.A., Wildung,M.R., Williams,D.C. and Croteau,R.B.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-2002) Institute of Biological Chemistry,
Washington State University, PO Box 646340, Pullman, WA 99164-6340,
USA
FEATURES
source
1. .2162
/organism="Pinus taeda"
/mol_type="mRNA"
/db_xref="taxon:3352"
59..1942
/notes="monoterpene synthase; Pt10"
/codon_start=1
/product="alpha-terpineol synthase"
/protein_id="AA061227.1"
/db_xref="GI:28894486"
/translation="MDLISVLPSASKSVCLHKPLSSSTHKLKPKFKTIRILVMPRRW
ERAPMSJSTVASEDDIQRRTGGYLSNLWDDVIOFLSTPYGELAYRERAEILIDEV
RIFSSMSDELGFSLIQLRMVNDVADLGRFKNIKSDALDYVYSEKLGIGC
GKSIITNLNSALGFRLLRWGPVSADVLKHFNRQIQGVFSCPSETBEDIRIMVNL
YRSLAFVAPGKVEEASFSEKYLKTLQKIDCSLSREIGDVLHGWHNLP
RLEARNVIDFGDDTKMEPNRKTLELAKLEFNI FOSIOKTELESLLRWNDGSG
POITFRHREYTYLASCIAPQHSGRFGFAKCHILTVLDDMIDLFQVDEKL
FTALIKRWDPSATDCLPQMKGLYIMVINTVMSABAKQAGRDTLNVARQWEDCL
DSHQAKWATGFLTFEYLENGKVSAAHRSALQPLMTWDIPPPHILKEVDPS
NLNDLACMLRLRGDTRCYQADRARGEETSCISKMDNPGATEEDALNHLNVMISGV
IKELNWECLKPNSVPISSKINFDITRAFYGYKYRVDGYSVSVETKSLVMRTLEP
VPL"
ORIGIN
Alignment Scores:
Pred. No.: 4,1e-175 Length: 2162
Score: 2266.50 Matches: 443
Percent Similarity: 82.31% Conservative: 78
Best Local Similarity: 69.98% Mismatches: 91
Query Match: 69.72% Indels: 21
DB: 8 Gaps: 9
US-10-025-145A-65 (1-618) x AF543529 (1-2162)
QY 1 MetAlaLeuSerIleThrProLeuValSerArgSerCysLeu----- 15
Db 59 ATGGATTATATCTCTACCGTCGGCTTCCAAATCGTGTGTGCTGCACAAACCC 118
|||||
QY 16 -----SerSerSerHisGluIleLysAlaLeuArgArgThrIleProThrLeuGlyIle 33
|||||
Db 119 TTGAGTAGTCTCTACTCATAACTTAAGCCTTTCTGCACAAACCAATCGGATTCTTTGAATG 178
|||||
QY 34 CysArgProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSer 53
|||||
Db 179 CCGAGGCGATGGGAATTTGCCAGCCCTCC-----ATGAGTCTGAGTACCGTTCATCT 232
|||||
QY 54 ThrAspSerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPhe 73
|||||
Db 233 GAAGATGATATACAAAGACGACCGCGGTATCTTTCCAACTGTGGAACGATGATGTG 292
|||||
QY 74 IleGlnSerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluAraGalaAspArg 93
|||||
Db 293 ATACAG---TTTCTGTCAACGCCTTATGGGAACTCGCTTACCGTGAACGTGCTGACCGA 349
|||||
QY 94 LeuIleGlyGluValLysAspIleMetPheAsnPhelLysSerLeuGluAspGlyGly--- 112
|||||
Db 350 CTGATTGATGAAGTAAAGGACATA---TTGAGTTCGATGTCATTTGGAGATGAGAAATTC 406
|||||
QY 113 AsnAspLeuLeuGlnArgLeuLeuValAspValGluArgLeuGlyIleAspArg 132
|||||
Db 407 AGTGATCTCATCAACGCCTTTGGATGGTCGATAAAGCTTGAAGCTTTGGGATCGATAGG 466
|||||
QY 133 HisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsnGluLys 152
|||||
Db 467 CAITTCAAAATGAGATCAATCAGCGCTTGATTTGTTTACAGTTACTCGAGCGAAAAA 526
|||||
QY 153 GlyIleGlyCysGlyArgGluSerValThrAspLeuAsnSerThrAlaLeuGlyLeu 172
|||||
Db 527 GGCATTGGATGTGGACAAAAAGTATTATTACTTAATCTCAATTCATCTGGCGCTT 586
|||||
QY 173 ArgThrIleuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPheLysAsp 192
|||||
Db 587 CGAACTCTTCGATTACACGATACCCCTGTTCTTCGACAGCGTTTGAACATTTTCAGAAAC 646
|||||
QY 193 LysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGluIleArgGlyVal 212
|||||
Db 647 CAAATTGGGCAGTTTGTATCATCTCTCTAGT---GAGACAGAGGAAGACATPAAGATCATG 703
|||||
QY 213 LeuAsnLeuPheArgAlaSerLeu-----ValalaPheProGlyGluLysVal 228
|||||
Db 704 GTCAATTTATATCGGCTCTCCCTCATTCGCTTTCCTGCTTCCCGTGGAGAAAGTT 763
|||||
QY 229 MetAspGluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIlePro 248
|||||
Db 764 ATGGAGAGCGAGAAAGCTTCTCAGAGAAATATTTAAAGGAAACTCTGCAAAAGATTCCG 823
|||||
QY 249 AlaSerSerIleLeuSerLeuGluIleAlaGspValLeuGluTyrGlyTyrHisThrAsn 268
|||||
Db 824 GACTGCAGT---CTTTCAAGAGAGATAGGGACGTTTGGACRTGTTGGCACACAAAT 880
|||||
QY 269 LeuProArgLeuGluAlaAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsn--- 287
|||||
Db 881 TTGCCCGGATTAGAACCAAGGAATTTACATCGACGCTCTTCGACACAGACACTAAGAAATATG 940
|||||
QY 288 -----LysAsnAlaAlaGluLysLeuLeuAlaLysLeuLeuAlaLysLeuPheAsnIlePhe 305
|||||
Db 941 GAGCCAAATAGGAAGACGGAGAAACTGTAGAACTTGCAGAACTGGAGATTCACATCTTT 1000
|||||
QY 306 HisSerLeuGlnGluArgGluLeuLysHisValSerArgTrpTrpLysAspSerGlySer 325
|||||
Db 1001 CAATCCATACAGAGACAGATTAGAGTCCCTCTTCAGATGGTGGATGATTTCAGGTTTCG 1060
|||||
QY 326 ProGluMetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIle 345
|||||
Db 1061 CCTCAGATTACCTTCATCGACATCGCCACGTTGGAGTACTACACTTTGGCTTCTCTGCATC 1120
|||||
QY 346 AlaPheGluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIle 365
|||||
Db 1121 GGGTTTGAGCTTCAACATTTCTGGATTACAGCTCGGCTTTCGGAAGCTTGTATATCTC 1180
|||||
QY 366 ThrValLeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThr 385
|||||
Db 1181 ACTGTTCTCGAGATGATGTACGACCTCTTTCGGAACAGTTTCGAGCTCAAACTTTTCACA 1240
|||||
QY 386 AlaThrIleLysArgTrpAspProSerAlaMetGluCysLeuProGluTyrMetLysGly 405
|||||
```

```
Db 1241 GCCGCAATTAGAGATGGATCCGTCGCCACAGATTGCCCTCCACAATATATGAAGGA 1300
Qy 406 ValTyrMetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGln 425
Db 1301 ATTTACATGATGGTTTACACACCGTAATGAATGTCTCGGAGGCACAGAAGGCTCAA 1360
Qy 426 GlyArgSerThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMet 445
Db 1361 GCCGAGACATCTCAACTATGCTCGACAGGCTTGGAGAGATTGTCTTGATTCACATAG 1420
Qy 446 GlnGluAlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsn 465
Db 1421 CAAGAGCAAGTGGATCGCCACCTGTTTCTGCCAACATTCGAGGAATATCTTGAGAAC 1480
Qy 466 GlyLysValSerSerAlaHisArgProCysAlaLeuGlnProfileLeuThrLeuAspile 485
Db 1481 GGGAAAGTTAGCTCTGCTCATCGCGTATCGGCATTTGCCAACCCATGCTGACGATGGACATC 1540
Qy 486 ProPheProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspile 505
Db 1541 CCCTTCCCTCTCATCTCTCAAGGAAGTTGACTTTCATCCAACTTCCATGACTTGGCA 1600
Qy 506 CysIleIleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGly 525
Db 1601 TGTCCCATGCTTCGATTACGAGGTGATACCCGATGTTATCAGCGGCACAGGCTCGTGG 1660
Qy 526 GluGluAlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAsp 545
Db 1661 GAAGAAACCTCTCTATATCTTGTATATGAAGACAATCTCGGAGCAACAGAGAGAT 1720
Qy 546 AlaLeuAsnHisIleAsnPheIleArgAspAlaIleArgGluLeuAsnTrpGluLeu 565
Db 1721 GCTCTTATCACTCAACGTCATGATCAGTGGCGTAATTAAGGAATTAATTTGACATTACCGA 1780
Qy 566 LeuLysProAspAsnSerValProIleThrSerLysLysHisAlaPheAspileSerArg 585
Db 1781 CTCAACCCACACAGAGTGTGCCCATTTCTTCCAGAAAAATTAATTTGACATTACCGA 1840
Qy 586 ValTrpHisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLys 605
Db 1841 GCCTTCCATTACGGTTACAAATACCGAGATGGCTACACGCTTCCAGCGTTGAAACAAAG 1900
Qy 606 SerLeuValMetArgThrValIleGluProValProLeu 618
Db 1901 AGTTTGTGTATGAGAACCTCTTGAACCTGTGCTTTA 1939

RESULT 14
AR222137 2186 bp DNA linear PAT 26-SEP-2002
LOCUS
DEFINITION Sequence 66 from patent US 6429014.
ACCESSION AR222137
VERSION AR222137.1 GI:23329511
KEYWORDS
SOURCE
ORGANISM
Unclassified.
REFERENCE 1 (bases 1 to 2186)
AUTHORS Steele,C.L., Bohlmann,J. and Croteau,R.B.
TITLE Monoterpene synthases from grand fir (Abies grandis)
JOURNAL Patent: US 6429014-A 66 06-AUG-2002;
FEATURES
    Location/Qualifiers
        source
            1..2186
                /organism="unknown"
                /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 6.65e-175 Length: 2186
Score: 2264.00 Matches: 435
Percent Similarity: 80.53% Conservative: 78
Best Local Similarity: 68.29% Mismatches: 98
Query Match: 69.64% Indels: 26
DB: 6 Gaps: 9
```

US-10-025-145A-65 (1-618) x AR222137 (1-2186)

```
Qy 1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15
Db 34 ATGGCTCTGGTTTCTTCGGACCC-----AAATCCTCGCTGCACAAATCGTTGATC 84
Qy 16 ---SerSerSerHisGluIleLysAlaLeuArgThrIleProThrLeuGlyIleCys 34
Db 85 AGGTCTACTCATCATGAGCTCAAGCCTCTCGCAGAACCATCCCAACTCTTGGAAATGFT 144
Qy 35 ArgProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThr 54
Db 145 AGGCGAGGGAATCTTTTCACACCTTCTGTGACATGATTTTGACCAACCGCTGTCTCTGAT 204
Qy 55 AspSerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIle 74
Db 205 GATGGTCTACAAAGACGATAGGTGACATCATTTCCAAATCTCTGGGACGAGATTTCATA 264
Qy 75 GlnSerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeu 94
Db 265 CAGTCTCTA---TCAACGCGCTTATGGGAGCCTTCTTACCAGAGAACGTCGTGAGAACTG 321
Qy 95 IleGlyGluValLysAspIleMetPheAsnPhelysSerLeuGluAspGlyGly----- 112
Db 322 ATTCGGGAAAGTGAAGGAG---ATGTTCAATTCATCGCATCGAAGATGGAGAAATCAATG 378
Qy 113 -----AsnAspLeuGlnArgLeuLeuValAspValGluArgLeuGly 129
Db 379 AGTCCCTCAATGATCTTATGAACGACTTGGATGGTCGATAGCGTTGAACGTTGGGG 438
Qy 130 IleAspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrp 149
Db 439 ATTCATAGACATTTCAAAAAAGACATAAAATCAGCCCTTGATTAAGTTACAGTTATTGG 498
Qy 150 AsnGluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAla 169
Db 499 AACCAAAAAGGTATGTGATCGGTAGAGATAGTGTCTTCTCAGAGGTTTGAAGATA 558
Qy 170 LeuGlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnVal 189
Db 559 TCGGGGTTTGAACCTCTTCGCTACCGATACAGTGTCTTTCAGAGGTTTGAAGATA 618
Qy 190 PheLysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGluIle 209
Db 619 TTTCAAGACCAAAATGGCAGTTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 675
Qy 210 ArgGlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMet 229
Db 676 AGAACCGTTCTGAATTTATATCGGGCTTCTTTCATTCGCTTCTCTCGGGGAGAAATGATG 735
Qy 230 AspGluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAla 249
Db 736 GAAGAGGCTGAATTTTCTCTCAAGATATTTGAAAGAGCCGTGCAAAAGATTCGGGTC 795
Qy 250 SerSerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTyrHisThrAsnLeu 269
Db 796 TCCAGT---CTTTCAAGAAATAGACATACATTTTGAATATGTTGGCACACAAATATG 852
Qy 270 ProArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThr----- 285
Db 853 CCAAGATTGAAACACAGGAATTAATTAGATGTTATTTGGACATCTCTACAGTCCATGGCTC 912
Qy 286 LysAsnLysAsnAla-----AlaGluLysLeuLeuAlaLysLeuGlu 301
Db 913 AAGAGAAAGGACGCAATATCTGGACAGCGAAAGCTTTTAGAATTCGCAAAATTCGAG 972
Qy 302 PheAsnIlePheHisSerLeuGlnArgLubLeuLysHisValSerA-gTTrpTrpLys 321
Db 973 TTCAACATCTTTCATCTCCCTTCAACAGAGGAGTTACAGTATCTCTCCAGATGGTGGATA 1032
Qy 322 AspSerGlySerProGluMetThrPheCysArgHisArgHisValGluTyrTyrAlaLeu 341
```

1033 CATTGGGTTTGCCTGAACCTGACCTTTGGTCGCATCGTCACGTGGAAATACATACACCCCTG 1092

342 AlaSerCysIleAlaPheGluProGlnHisSerGlyPheArgLeuGlyPheThrIysMet 361

1093 AGCTCTTGCAATGCGACTGAGCCCAACAACTTCGATTCAGATGGCTTGCCTCAAAACG 1152

362 SerHisLeuIleThrValLeuAspMetTyrAspValPheGlyThrValAspGluLeu 381

1153 TGTCACTTATACGGTCTTGACGATATCTACGACACTTTCGACAGATGATGAATC 1212

382 GluLeuPheThrAlaThrIleLysArgTrpAspProSerAlaMetGluCysLeuProGlu 401

1213 GAACTCTTCAACGAGGCACTTAGAGATCGAATCCGTCGAGAAAGACGCTCCAGAA 1272

402 TyrMetLysGlyValTyrMetMetValTyrHisThrValAsnGluMetAlaArgValAla 421

1273 TATATGAAGAATCTACATGCACTCTACGAAGCCTTAACGACATGGCGGAGAGGCA 1332

422 GluLysAlaGlnGlyArgAspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPhe 441

1333 GAGAAGACACAGGCGGACACGCTCAATATGCTAGAAAGCTTGGAGTTTATCTT 1392

442 AspSerTyrMetGlnGluAlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGlu 461

1393 GATTCGTATACACAAAGCAAGTGAATCGCCAGCGTTATCTGCCAACTTTCGAGGAG 1452

462 TyrLeuGluAsnGlyLysValSerSerAlaHisArgProCysAlaLeuGlnProIleLeu 481

1453 TACTTAGAAGCGGAGGTTAGCTCTGTCATCTGTCGACGCGCATGTACACCCCTCTG 1512

482 ThrLeuAspIleProPheProAspHisIleLeuLysGluValAspPheProSerLysLeu 501

1513 ACATTGACGCTACCGCTTCTGATGACGCTTGAAGGGAATAGATTTTCATCGAGATT 1572

502 AsnAspLeuIleCysIleIleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAsp 521

1573 AATTGGGAATCTCTCAACCCGATAGCAATATTCGAATGACTGACGGAACATGCTTAT 1612

522 ArgAlaArgGlyGluAlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeu 541

1633 AGGACCGAGAGGAGGAGCGTCAAGCATGCTGTTACATGAAGCAATCCCGGATTA 1692

542 ThrGluGluAspAlaLeuAsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeu 561

1693 ACAGAGGAGATGCTCTCATCATATCATGTCATGATCAACGACATATCAAGAAATTA 1752

562 AsnTrpGluLeuLysProAsnSerValProIleThrSerLysIysHisAlaPhe 581

1753 AATTGGGAATCTCTCAACCCGATAGCAATATTCGAATGACTGACGGAACATGCTTAT 1812

582 AspIleSerArgValTrpHisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsn 601

1813 GAGATCAACGAGCTTTCCACCACTTTTCAATATAGATAGCTTTCAGCGTTGCCACT 1872

602 ValGluThrLysSerLeuValMetArgThrValIleGluProValProLeu 618

1873 CAAGAAACGAAAGTTTGGTAGGAGACGGTCTTGAACCATGCTCTCT 1923

RESULT 15

AF139205

LOCUS

DEFINITION

ABIES GRANDIS beta-phellandrene synthase (agc8) mRNA, complete cds.

AF139205

VERSION

AF139205.1 GI:7381248

KEYWORDS

SOURCE

ORGANISM

Abies grandis

Abies grandis

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.

REFERENCE

1 (bases 1 to 2186)

AUTHORS

Bohlmann, J., Phillips, M., Ramachandiran, V., Katoh, S. and Croteau, R.

TITLE

cDNA cloning, characterization, and functional expression of four new monoterpene synthase members of the tpsd gene family from grand

fir (Abies grandis)

Arch. Biochem. Biophys. 368 (2), 232-243 (1999)

99373092

10441373

2 (bases 1 to 2186)

Phillips, M.

Direct Submission

Submitted (30-MAR-1999) Institute of Biological Chemistry,

Washington State University, 289 Clark Hall, Pullman, WA

99164-6340, USA

Location/Qualifiers

1..2186

/organism="Abies grandis"

/mol\_type="mRNA"

/db\_xref="taxon:46611"

/tissue\_type="stem"

/dev\_stage="sapling"

1..2186

/gene="agc8"

/note="TpsD gene family member"

34..1926

/gene="agc8"

/note="monoterpene synthase; includes N-terminal transit peptide"

/codon\_start=1

/product="beta-phellandrene synthase"

/protein\_id="AAF61453.1"

/db\_xref="GI:7381249"

/translation="MALVSSAPKSLHSLRSTHHELPKLRPTIPTLMCRGKSFT

PSVMSLTAVSDGDLQRIGDYHNSLMDDDFIQSLSTPYGSPSFRERAEKLIQVEKE

MENSPSEDSGMSPLNDLIERLMGVDSVERLIGDRHFKKEIKSALDYVYSYWNKGI

CGCRDVSFVDNASTAGFTLRLHGVSVSEVLKVFQDQNGQAFSPSKERDIRTVL

NLYRASFIAPGCKVMEBAEIFFSSRYLKEAVOKIPVSSLSQIDYVLEYGWTNMPRL

ETRNLYDPGHTSPMLKKPTQYLDSEKLLEKLEFNIHSLQKELQYLSRWMIH

SGELPTGRHRHVEYTYLSSCIATEPHSAFLGFAKTCHLITVLDLDYDFGTMDE

ILDFNEAVRWMPSEKERLPEYKEIYMALEYALTDMAREAKTQGRDINTFARKWE

VLYDTPQAKMIAAGYLPTPEEYENAKVSSGHRAAALTPLLTLDVDPDVLKGI

FFSRFNDLASSFLRGDTRCYKADRDRGEASISCYMKDNPGLTEEDALHNAMI

NDIIEKLNWELLKPSDNIPTARKHAYEITRAFHLQYKYRDRGFSVATQKSLVRRTV

LEPVL"

ORIGIN

Alignment Scores:

Pred. No.: 6,65e-175 Length: 2186

Score: 2264.00 Matches: 435

Percent Similarity: 80.53% Conservativeness: 78

Best Local Similarity: 68.29% Mismatches: 98

Query Match: 69.64% Indels: 26

DB: 8 Gaps: 9

US-10-025-145A-65 (1-618) x AF139205 (1-2186)

QY 1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15

DB 34 ATGGCTCTGGTTCTTCCGACCC-----AAATCTGCTGCTGCACAAATCGTTGATC 84

QY 16 ---SerSerHisGluIleLysAlaLeuArgArgThrIleProThrLeuGlyIleCys 34

DB 85 AGGTCTACTCATCATGAGTCAAGCTCTGCGCAGAACCATCCCACTCTTGAATGTGT 144

QY 35 ArgProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThr 54

DB 145 AGGCGAGGGAATCTTTCACACCTCTGTCGATGAGTTTGACCCCGCTGATCTGAT 204

QY 55 AspSerValGluArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIle 74

DB 205 GATGGTCTCAAGACGATAGTACTATTCATTCATCTCTGGACACGATTTTCATA 264

QY 75 GlnSerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeu 94

DB 265 CAGTCTCTA--TCAACGCTTATGGGAGCTTCTTACCAGAACGCTGCTGAGAACTG 321

QY 95 IleGlyGluValLysAspIleMetPheAsnPheLysSerLeuGluAspGlyGly----- 112

```
Db 322 ATTGGGGAAGTGAAGGAG---ATGTTCAATTCAATGCCATCGGAAGATGGAGAAATCAATG 378
Qy 113 -----AsnAspLeuGlnArgLeuLeuValAspAspValGluArgLeuGly 129
Db 379 AGTCCCTCAATGATCTATTGAACGACTTTGGATGGTTCGATACGTTGAACGTTTGGGG 438
Qy 130 IleAspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrp 149
Db 439 ATTGATAGACATTTCAAAAAAGAGATAAAATCAGCCCTTGATTATGTTTACAGTTATTGG 498
Qy 150 AsnGluLysGlyIleGlyCysGlyArgGluSerValThrAspLeuAsnSerThrAla 169
Db 499 AACGAAAAAGGTATTGGATGGGTAGAGATAGTGTTCCTGATGTCACACTCAACTGCACTGCC 558
Qy 170 LeuGlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnVal 189
Db 559 TCGGGGTTTCGAATCTTCGGCTTACAGGATACAGTGTCTCTTCAGAGGTTTGAAGTA 618
Qy 190 PheLysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGluIle 209
Db 619 TTTCAGACCAAAATGGSCAGTTTGCATTCCTCTAGTACAAAA---GAGAGAGACATC 675
Qy 210 ArgGlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMet 229
Db 676 AGAACCCGTTCTGAATTTATATCGGGCTTCCTTCATTCCTGCTTCCTGGGGAGAAATGATG 735
Qy 230 AspGluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAla 249
Db 736 GAAGAGGCTGAAATTTCTCTCAAGATATTTTGAAGAGCGGTGCACAAAGATTCGGGTC 795
Qy 250 SerSerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTyrHisThrAsnLeu 269
Db 796 TCCAGT---CTTCAACAAGAAATAGACTACACTTTTGGAAATATGTTGGCACACAAATATG 852
Qy 270 ProArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThr----- 285
Db 853 CCAAGATTGGAACCAAGAAATTAATCTAGATGTTTGGACATCTCCACGATCCATGGGTC 912
Qy 286 LysAsnLysAsnAla-----AlaGluLysLeuLeuGluLeuAlaLysLeuGlu 301
Db 913 AAGAAGAAAGAGCGCAATATCTGGACAGCGAAAGCTTTTAGAATCTGCAAAATTTGGAG 972
Qy 302 PheAsnIlePheHisSerLeuGlnGluArgGluLeuLysHisValSerArgTrpTrpLys 321
Db 973 TTCAACATCTTCACTCCCTTCAACAGAGAGGATTACAGTATCTCTCCAGATGGTGGATA 1032
Qy 322 AspSerGlySerProGluMetThrPheCysArgHisArgHisValGluTyrTyrAlaLeu 341
Db 1033 CATTCGGGTTTGCCTGAACTGACCTTGGTGGCATCGTCAGTGGAAATCTACACCCCTG 1092
Qy 342 AlaSerCysIleAlaPheGluProGlnHisSerGlyPheArgLeuGlyPheThrLysMet 361
Db 1093 AGCTTTGCAATTGCACTGAGCCCAACATCTGCATTGAGTTGGGCTTTGCCAAAACG 1152
Qy 362 SerHisLeuIleThrValLeuAspAspMetTyrAspValPheGlyThrValAspGluLeu 381
Db 1153 TGTCACTTTATCAGCGTTCTGGACGATATCTACGACACTTTCGGAACGATGGATGAAATC 1212
Qy 382 GluLeuPheThrAlaThrIleLysArgTrpAspProSerAlaMetGluCysLeuProGlu 401
Db 1213 GAACTCTTCAACGAGCGAGTTAGGAGATGGAATCCGTGCGAGAAAGACGCTCCAGAA 1272
Qy 402 TyrMetLysGlyValTyrMetMetValTyrHisThrValAsnGluMetAlaArgValAla 421
Db 1273 TATATGAAAGAATCTACATGGCACTCTACGAGCGCTTAACCTGACATGGCGGAGAGGCA 1332
Qy 422 GluLysAlaGlnGlyArgAspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPhe 441
Db 1333 GAGAAGACACAAGGCGGACACGCTCAATTTATGCTAGAAAGGCTTGGGAAGTTTATCTT 1392
Qy 442 AspSerTyrMetGlnGluAlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGlu 461
```

```
Db 1393 GATTCGTATACACAGAAAGCAAAAGTGGATGCCAGCGGTTATCTGCCAACTTTTCAGGAG 1452
Qy 462 TyrLeuGluAsnGlyLysValSerSerAlaHisArgProCysAlaLeuGlnProIleLeu 481
Db 1453 TACTTAGAGAAACGGAAGGTTAGCTCTGGTTCATCGTCAGCGGCATTGACACCCCTCTG 1512
Qy 482 ThrLeuAspIleProPheProAspHisIleLeuLysGluValAspPheProSerLysLeu 501
Db 1513 ACATTGGACGTACCGCTTCCTGATGACGTCTTGAAGGGAATAGATTTTCCATCGAGATTT 1572
Qy 502 AsnAspLeuIleCysIleIleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAsp 521
Db 1573 AATGATTTGGCATCTTCCTTCCTTAGACTAAGAGTGCACACACGATGCTACAGGCGAC 1632
Qy 522 ArgAlaArgGlyGluGluAlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeu 541
Db 1633 AGGACCCGAGGAGAGAAAGCGTCAAGCATATCGTGTTCATGAAAGACAATCCCGGATTA 1692
Qy 542 ThrGluGluAspAlaLeuAsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeu 561
Db 1693 ACAGAGGAAGATGCTCTCAATCATATCAATGCCCATGATCAACGACATAATCAAGAAATTA 1752
Qy 562 AsnTrpGluLeuLeuLysProAspAsnSerValProIleThrSerLysLysHisAlaPhe 581
Db 1753 AATTGGGAACCTTCTCAACCCATAGCAATATTCAATGACTGCACGGAACATGCTTAT 1812
Qy 582 AspIleSerArgValTrpHisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsn 601
Db 1813 GAGATAACGAGAGCTTTCCACCAACTTTACAAATATAGAGATGGCTTCAGCGTTGCCACT 1872
Qy 602 ValGluThrLysSerLeuValMetArgThrValIleGluProValProLeu 618
Db 1873 CAAGAAAACGAAAAGTTTGGTGAGGAGAACGGTCTCTTGAACCAAGTGCCTCTT 1923
```

Search completed: July 8, 2004, 19:20:23

Job time : 6830 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 8, 2004, 12:19:34 ; Search time 707 Seconds  
(without alignments)  
3713.418 Million cell updates/sec

Title: US-10-025-145A-65

Perfect score: 3251

Sequence: 1 MALLSITPLVSRSLSSSHE.....FANVETKSLVMRTVIEPVPL 618

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US10025145/runat\_07072004\_160149\_27925/app.query.fasta\_1.775  
-DB=N Geneseq\_29Jan04 -QPMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCI=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US10025145 @CGN 1.1 470 @runat\_07072004\_160149\_27925 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3251	100.0	2013	4 AAF73411	Aaf73411 Grand fir
2	2453	75.5	2018	2 AAX08644	Aax08644 Pinene sy
3	2453	75.5	2018	3 AAA38922	Aaa38922 Grand fir
4	2453	75.5	2018	4 AAF73372	Aaf73372 Grand fir
5	2264	69.6	2186	2 AAF73412	Aaf73412 Grand fir
6	2200.5	67.7	2196	2 AAX08643	Aax08643 Myrcene s
7	2200.5	67.7	2196	3 AAA38927	Aaa38927 Grand fir
8	2200.5	67.7	2196	4 AAF73371	Aaf73371 Grand fir

9	2197.5	67.6	2205	2	AAX08663	Aax08663 Grand fir
10	2197.5	67.6	2205	4	AAF73391	Aaf73391 Grand fir
11	2137	65.7	1890	4	AAF73421	Aaf73421 Grand fir
12	2084.5	64.1	2429	4	AAF73413	Aaf73413 Grand fir
13	2031.5	62.5	2089	2	AAX08645	Aax08645 Limonene
14	2031.5	62.5	2089	3	AAA38938	Aaa38938 Grand fir
15	2031.5	62.5	2089	4	AAF73373	Aaf73373 Grand fir
16	1617.5	49.8	1513	3	AAA69551	Aaa69551 Pinus rad
17	1611.5	49.6	1634	3	AAA69644	Aaa69644 Pinus rad
18	1295.5	39.8	1865	3	AAA38933	Aaa38933 Grand fir
19	1295.5	39.8	1885	3	AAA38933	Aaa38933 Delta-sel
20	1295.5	39.8	1885	2	AAX87534	Aax87534 Delta-sel
21	1295.5	39.8	1888	2	AAX87505	Aax87505 Grand fir
22	1291.5	39.7	1885	2	AAX87532	Aax87532 Delta-sel
23	1267.5	39.0	1173	3	AAA69643	Aaa69643 Pinus rad
24	1250.5	38.5	1967	2	AAX87513	Aax87513 Grand fir
25	1250.5	38.5	1967	2	AAX08655	Aax08655 Grand fir
26	1250.5	38.5	1967	4	AAF73383	Aaf73383 Grand fir
27	1240.5	38.2	2700	2	AAT97447	Aat97447 Pacific Y
28	1240.5	38.2	2700	3	AAA38931	Aaa38931 Yew taxad
29	1231	37.9	1785	2	AAX87535	Aax87535 Gamma-hum
30	1231	37.9	1785	3	AAA38934	Aaa38934 Grand fir
31	1231	37.9	1977	2	AAX87506	Aax87506 Grand fir
32	1230	37.8	1785	2	AAX87536	Aax87536 Gamma-hum
33	1230	37.8	1785	2	AAX87537	Aax87537 Gamma-hum
34	1188.5	36.6	2528	2	AAX87529	Aax87529 Grand fir
35	1187	36.5	2424	2	AAX08654	Aax08654 Grand fir
36	1187	36.5	2424	3	AAA38932	Aaa38932 Grand fir
37	1187	36.5	2424	4	AAF73382	Aaf73382 Grand fir
38	1187	36.5	2525	2	AAX87530	Aax87530 E-alpha-b
39	1187	36.5	2528	2	AAX87531	Aax87531 E-alpha-b
40	1187	36.5	2571	2	AAX87504	Aax87504 Grand fir
41	1052.5	32.4	2861	3	AAA38937	Aaa38937 Grand fir
42	856	26.3	696	4	AAF73414	Aaf73414 Grand fir
43	829.5	25.5	779	3	AAA69642	Aaa69642 Pinus rad
44	800	24.6	1416	2	AAX08656	Aax08656 Grand fir
45	800	24.6	1416	4	AAF73384	Aaf73384 Grand fir

#### ALIGNMENTS

RESULT 1  
AAF73411  
ID AAF73411 standard; cDNA; 2013 BP.  
XX  
AC AAF73411;  
XX  
DT 30-APR-2001 (first entry)  
XX  
DE Grand fir monoterpene synthase coding sequence fragment SEQ ID NO: 64.  
XX  
KW Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
KW terpinolene synthase; insect resistance; nutrition; ss.  
XX  
OS Abies grandis.  
XX  
PN WO200107565-A2.  
XX  
PD 01-FEB-2001.  
XX  
PF 24-JUL-2000; 2000WO-US020264.  
XX  
PR 26-JUL-1999; 99US-00360545.  
XX  
PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
XX  
PI Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
XX WPI; 2001-182782/18.  
XX P-PSDB; AAB69390.  
XX  
PT New nucleic acid encoding monoterpene synthases, for increasing terpene

PT synthesis in plants, e.g. for increasing resistance to pests or for  
 PT treatment of cancer.

XX Claim 8; Page 147-149; 175pp; English.

XX The present invention provides the protein and coding sequences of  
 CC monoterpene synthases from the grand fir. These include (-)-camphene  
 CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 CC limonene, (-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 CC and pinene synthase. The sequences can be used to produce transgenic  
 CC plants expressing high levels of the enzymes, resulting in levels which  
 CC are useful in protecting against and treating cancers, and to confer  
 CC insect resistance on plants

XX SQ Sequence 2013 BP; 591 A; 432 C; 464 G; 526 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 2013  
 Score: 3251.00 Matches: 618  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-025-145A-65 (1-618) x AAF73411 (1-2013)

QY 1 MetAlaLeuSerIleThrProLeuValSerArgSerCysLeuSerSerHisGlu 20  
 DB 36 ATGGCTCTCTTCTATTACTCGCTGGTTCAGGTGCGCTCAGTTCTTCATGAG 95  
 QY 21 IleLysAlaLeuArgThrIleProThrLeuGlyIleCysArgProGlyLysSerVal 40  
 DB 96 ATTAAGCTCTCGTAGAACATCCCACTCTTGGAACTGACGGCGGGGAATCCGTC 155  
 QY 41 AlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAspSerValGlnArg 60  
 DB 156 GCGCATTCATAAACATGTTGTGACAGCTGCGCATCTACTGATCTGTACAGAGCGC 215  
 QY 61 ValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIleGlnSerLeuIleSerThr 80  
 DB 216 GTGGGCAACTATCATCTCCAACTGTGGGACGATGATTTACAGTCTCTGATCAACG 275  
 QY 81 ProTyrGlyValaProAspTyrArgGluArgAlaAspArgLeuIleGlyGluValLysAsp 100  
 DB 276 CCTATGAGGACCTGATTAACGGGAACGTGCTGACAGACTTATTGGGAGTAAGGAT 335  
 QY 101 IleMetPheAsnPheLysSerLeuGluAspGlyGlyAsnAspLeuGlnArgLeuLeu 120  
 DB 336 ATAATGTTCAATTTCAAGTCGCTGGAAGATGGAGCAATGATCTCCTTCAACGACTTTG 395  
 QY 121 LeuValAspAspValGluArgLeuGlyIleAspArgHisPheLysLysGluIleLysThr 140  
 DB 396 GTGGTCGATGACGTTGAACTTTGGGAATCGACAGGCATTTCAAAAAGAGATAAACG 455  
 QY 141 AlaLeuAspTyrValAsnSerTyrTrpAsnGluLysGlyIleGlyCysGlyArgGluSer 160  
 DB 456 GCACCTGATTAATCTTAACAGTTATTGGAACGAAAGAGCATTTGGATGGGAGAGAT 515  
 QY 161 ValValThrAspLeuAsnSerThrAlaLeuGlyLeuArgThrLeuArgLeuHisGlyTyr 180  
 DB 516 GTTGTGACTGACCTCAACTCAACCGCTTTGGGCTTCGAACCTCCGACTACACGGATAC 575  
 QY 181 ThrValSerSerAspValLeuAsnValPheLysAspLysAsnGlyGlnPheSerSerThr 200  
 DB 576 ACTGTGCTTCAGATGTTTGAACGTTTTTAAAGACAAAATGGGCAATTTCTCCACT 635  
 QY 201 AlaAsnIleGlnLeuGluGlyIleArgGlyValLeuAsnLeuPheArgAlaSerLeu 220  
 DB 636 GCCAATATTACAGATAGAGGAGAGATTAGAGGCGTTCTCAATTTATTTCAGGGGCTCCCTC 695  
 QY 221 ValAlaPheProGlyGluLysValMetAspGluAlaGluThrPheSerThrLysTyrLeu 240  
 DB 696 GTCGCTTTCCCGCGGAGAAAGTTATGGATGAAGCTGAACATTTCTCTCAAAATATTTA 755

QY 241 ArgGluAlaLeuGlnLysIleProAlaSerSerIleLeuSerLeuGluIleArgAspVal 260  
 DB 756 AGAAGAGCCCTGCAAAAGATTCCGGCATCCAGTATACTTCTACTAGAGATACGGACGTT 815  
 QY 261 LeuGluTyrGlyTrpHisThrAsnLeuProArgLeuGluAlaArgAsnTyrMetAspVal 280  
 DB 816 CTGGATATGGTTGGCACACCAATTTGCCACGCTTGGAGCAAGGAATTTACATGGACGTC 875  
 QY 281 PheGlyGlnHisThrLysAsnLysAlaAlaGluLysLeuLeuLysLeuLeuLysLeu 300  
 DB 876 TTTGACAGCACACTAAAAATTAAGACCGCGCGAGAACTTTTAACTTGCATAATTG 935  
 QY 301 GluPheAsnIlePheHisSerLeuGlnArgGluLeuLysHisValSerArgTyrTrp 320  
 DB 936 GAATTCATATATTTTTCCTCTTACAGAGAGAGATTAAACATGTTTTCCCGATGGTGG 995  
 QY 321 LysAspSerGlySerProGluMetThrPheCysArgHisArgHisValGluTyrTyrAla 340  
 DB 996 AAAGACTCGGGTTCCTCTGAGATGACCTTCTGTGACATCGTCACGTGGAATACTACGCT 1055  
 QY 341 LeuAlaSerCysIleAlaPheGluProGlnHisSerGlyPheArgLeuGlyPheThrLys 360  
 DB 1056 TTGGCTTCTGCTTTCGAGCTCAACATCTCGAATTCAGACTCGGCTTTTACAAG 1115  
 QY 361 MetSerHisLeuIleThrValLeuAspAspMetTyrAspValPheGlyThrValAspGlu 380  
 DB 1116 ATGCTCATCTTATCAGGTTCTTGACGACATGTACACGCTTTCGGCAGCAGTAGACGAG 1175  
 QY 381 LeuGluLeuPheThrAlaThrIleLysArgTrpAspProSerAlaMetGluCysLeuPro 400  
 DB 1176 CTGGAACCTTTCACAGCGACAATTAAGAGATGGATCCGTCGCGATGGAATGCCCTTCCA 1235  
 QY 401 GluTyrMetLysGlyValTyrMetMetValTyrHisThrValAsnGluMetAlaArgVal 420  
 DB 1236 GAATATATCAAGAGAGTGTACATGATGTTTATCACCGTAAATGAATGGCTCCAGTG 1295  
 QY 421 AlaGluLysAlaGlnGlyArgAspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCys 440  
 DB 1296 GCAGAGAGGCTCAAGCCGACAGACAGCTCACTATGCAAGACAGCGCTTGGGAGGCGTGT 1355  
 QY 441 PheAspSerTyrMetGlnGluAlaLysTrpIleAlaThrGlyTyrLeuProThrPheGlu 460  
 DB 1356 TTGATTCGTATATGAGGAAGCAAGTGGATCGGCTGCTTATCGCCACGTTTGTAG 1415  
 QY 461 GluTyrLeuGluAsnGlyLysValSerSerAlaHisArgProCysAlaLeuGlnProIle 480  
 DB 1416 GAGTACTTGGAGAACGGGAAAGTTAGCTCTGCTCATCGCCCATGCGCAGCTGCAACCCAT 1475  
 QY 481 LeuThrLeuAspIleProPheProAspHisIleLeuLysGluValAspPheProSerLys 500  
 DB 1476 CTGACGTTGGACATCCCTTTCCTGATCATCTTCAAGGAAGTTGACTTCCCATCGAAG 1535  
 QY 501 LeuAsnAspLeuIleCysIleIleLeuArgLeuArgGlyAspThrArgCysTyrLysAla 520  
 DB 1536 CTCAATGACTTGATATGATCATCTTCGATTAAGAGGTGATACACGCTGCTACAGGCA 1595  
 QY 521 AspArgAlaArgGlyGluGluAlaSerSerIleSerCysTyrMetLysAspAsnProGly 540  
 DB 1596 GACAGGCGCGTGGAGAAAGCTTCGCTCTATCATGTTTATGAAAGCAATCTCTGGA 1655  
 QY 541 LeuThrGluGluAspAlaLeuAsnHisIleAsnPheMetIleArgAspAlaIleArgGlu 560  
 DB 1656 TTAACGGAAGAGATGCTCTGAATCATATCACTTCATGATCAGGAGCGCAATCAGAGAA 1715  
 QY 561 LeuAsnTrpGluLeuLeuLysProAspAsnSerValProIleThrSerLysLysHisAla 580  
 DB 1716 TTAATTTGGAGCTTCTAAAGCCAGACACAGTGTTCCTCATCTTCCCAAGAAACACGCA 1775  
 QY 581 PheAspIleSerArgValTrpHisHisGlyTyrArgTyrArgAspGlyTyrSerPheAla 600  
 DB 1776 TTTGACATTAAGCAGAGTTTGGCATCACGGTTACAGATACCGAGATGCGTACAGCTTGGC 1835





246	Db	TCTTTA---CCAA	CGCGTTATGAGGAAAAATCGTACTCGGAGCGTGTGTGAGAAACTGATC	302
96	Qy	GlyGluValIysAspIleMetPheAsnPheIysSerLeuGluAspGlyGly	-----	112
303	Db	GGGGAAGTAAGAAC--	ATGTTCAATTCGATGTCATTAGAGATGAGAGTTAATGAGT	359
113	Qy	-----AsnAspLeuLeuGlnArgLeuLeuLeuValAspAspValGluArgLeuGlyIle	130	
360	Db	CCGCTCAATGATCTCAVTC	CAACGCGTTTGGATTCTGCACAGCCCTGAACGTTTGGGATC	419
131	Qy	AspArgHisPheIysLysGluIleIysThrAlaLeuAspTyrValAsnSerTyrTrpAsn	150	
420	Db	CATAGACATTTCAAGATGAGATAAAATCGCGCTTGATTA	GTGTTACAGTTATTTGGGCGC	479
151	Qy	GluIysGlyIleGlyCysGlyValArgGluSerValValThrAspLeuAsnSerThrAlaLeu	170	
480	Db	GAAATGGCATCGATCGGAGGAGAGTGTGTTACTGATCTGA	CTGAACCTCACTGCGGTG	539
171	Qy	GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe	199	
540	Db	GGGCTTCGAACCTTACGACTACACGGATACCCGGTGTCTT	CAGATGTTTTCAAGCTTTC	599
191	Qy	LysAspIysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGluIleArg	210	
600	Db	AAAGCCAAAATGGCGAGTTTTCTCGTCTGAAATATTCAGACAGATGAGAGATCAGA	659	
211	Qy	GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluIysValMetAsp	230	
660	Db	GGCGTCTTGANTTATTTCCGGGCTCCCTCATTCGCCCTTCCAGGGGAGAAAAAT	TATGGAT	719
231	Qy	GluAlaGluThrPheSerThrLysTyrIleuArgGluAlaLeuGlnIysIleProAlaSer	250	
720	Db	GAGGCTGAAATCTTCTCTACCAATATTTAAAGAGAGCCCTGCAAAAGATTC	CGGTCTCC	779
251	Qy	SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTyrHisThrAsnLeuPro	270	
780	Db	AGT---CTTTTCGCGAGAGATCGGGAGCGTTTGGAAATATGGTGGCACACATAT	TTTCGCG	836
271	Qy	ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsnLysAsnAla	290	
837	Db	CGATTGGAAGCAAGGAATTACATCCAAGTCTTTTGGACAGGACACTGAGACACGAAGTCA	896	
291	Qy	-----AlaGluLysLeuLeuGluLeuAluLysLeuGluPheAsnIlePheHisSer	307	
897	Db	TATGTGAAGAGCAAAAAAATTTTAGAACTCGCAAAATTTGAGTTTCAACATCTT	CAATCC	956
308	Qy	LeuGlnGluArgGluLeuLysHisValSerArgTrpTrpIysAspSerGlySerProGlu	327	
957	Db	TTACAAAGAGGGAGTTAGAAAATTTTCGGTCAGATGGTGGAAAGAAATCGGGT	TTTCCGTGAG	1016
328	Qy	MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe	347	
1017	Db	ATGACCTTCTGCCAATCTGTCAGCTGGATACACTTACCTTTGGCTTCTCGATTGCGTTC	1076	
348	Qy	GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal	367	
1077	Db	GAGCCTCAACATCTCGGATTCCAGACTCGGCTTCCCAAGAGCTGTCATCTTATC	CACGGTT	1136
368	Qy	LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluPheThrAlaThr	387	
1137	Db	CTTGACCATATGTACACACCTTCGGCACAGTACAGCAGCTGGAACTCTTTCACACG	CAGCA	1196
388	Qy	IleLysArgTrpAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr	407	
1197	Db	ATGAAGAGATGGGATTCGCTCTCGATGATTTGCTTCCAGAAATATATGAAGGAGTGTAC	1256	
408	Qy	MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg	427	
1257	Db	ATAGCGGTTACGACACCGTAAATGAAATGCGCTCGAGGGCAGAGGAGGCTCAAGG	CGCA	1316
428	Qy	AspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGlu	447	
1317	Db	GATACGCTCATAATGCTCGGAAGCTTGGAGGCTTATATGATTCGTATATCAAGAA	1376	

Qy	448	AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAenGlyLys	467
Db	1377	GCAAGTGGATCGCCACTGGTTACTGCCCTCCTTTGATGACTACGAGAATGGAAA	1436
Qy	468	ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe	487
Db	1437	GTTAGCTGTGGTCATCCGATATCCGCAATTCGACCCCAATTCGACATGGACATCCCCCTTT	1496
Qy	488	ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle	507
Db	1497	CCTGATCATATCCTCAAGGAAGTTGACTTCCCATCAAAGCTTAAACGACTTGGCATGTGCC	1556
Qy	508	IleLeuArgLeuArgGlyAspThrArgCysTyrIlysAlaAspArgAlaArgGlyGluGlu	527
Db	1557	ATCCTTCGATTACAGGTGATACGCGTGTCTACAAGGCGGACAGGGCTCGTGGAGAAGAA	1616
Qy	528	AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu	547
Db	1617	GCITTCCTCTATATCATGTTATATGAAGACAATCCTCGAGTATCAGAGGAAGATGCTCTC	1676
Qy	548	AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLys	567
Db	1677	GATCATATCAAGCGCCATGATCAGTGCATTAACAAGGATTAATTTGGAACTTCACAA	1736
Qy	568	ProAspAsnSerValProIleThrSerLysIlyHisAlaPheAspIleSerArgValTrp	587
Db	1737	CCAGACATCAATGTTCCCATCTCGCGGAAGAAAACATGCTTTTGACATCGCCAGAGCTTTC	1796
Qy	588	HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu	607
Db	1797	CATTACGGCTACAATAACGAGACGGCTACAGGTTGCCAACGTTGAAACGAAAGATTTC	1856
Qy	608	ValMetArgThrValIleGluProValProLeu	618
Db	1857	GTCACGAGAACCCCTCCTTGAATCTGTGCCTTTG	1889
RESULT 4			
AAAF73372			
ID	AAAF73372	standard; cDNA; 2018 BP.	
XX	AC	AAAF73372;	
XX	AC	AAAF73372;	
XX	DT	30-APR-2001 (first entry)	
XX	DE	Grand fir (-)-pinene synthase coding sequence SEQ ID NO: 3.	
XX	KW	Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;	
KW	KW	myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;	
KW	KW	terpinolene synthase; insect resistance; nutrition; ss.	
XX	OS	Abies grandis.	
XX	XX	WO200107565-A2.	
XX	XX	01-FEB-2001.	
XX	XX	24-JUL-2000; 2000WO-US020264.	
XX	XX	26-JUL-1999; 99US-00360545.	
XX	XX	(UNIW ) UNIV WASHINGTON STATE RES FOUND.	
XX	PA	Steele CL, Bohlmann J, Croteau RB, Phillips MA;	
XX	PI	WPI; 2001-182782/18.	
XX	DR	P-PSDB; AAB69371.	
XX	XX	New nucleic acid encoding monoterpene synthases, for increasing terpene	
PT	PT	synthesis in plants, e.g. for increasing resistance to pests or for	
PT	PT	treatment of cancer.	
XX	XX	Claim 38; Page 108-110; 175pp; English.	

XX

The present invention provides the protein and coding sequences of monoterpene synthases from the grand fir. These include (-)-camphene synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase and pinene synthase. The sequences can be used to produce transgenic plants expressing high levels of the enzymes, resulting in levels which are useful in protecting against and treating cancers, and to confer insect resistance on plants

SQ Sequence 2018 BP; 583 A; 431 C; 479 G; 525 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	1,2e-243	Length:	2018
Score:	2453.00	Matches:	475
Percent Similarity:	84.63%	Conservative:	59
Best Local Similarity:	75.28%	Mismatches:	81
Query Match:	75.45%	Indels:	16
DB:	4	Gaps:	6

US-10-025-145A-65 (1-618) x RAF73372 (1-2018)

QY	1	MetAlaLeuSerIleThrProLeuValSerArgSerCysLeu-----	15
DB	6	ATGGCTAGTTTCTACCGCACCCTGGCTTCCAAATCATGCTGCACAAATCGTTGATC	65
QY	16	SerSerSerHisGluIleLysAlaLeuArgArgThrIleProThrIleGlyCysArg	35
DB	66	AGTTCTACCCATGAGCTTAAGGCTCTCTAGAACAAATCCAGCTCTAGAAATGATAGG	125
QY	36	ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp	55
DB	126	CGAGGGAATCTATCACTCTCCATCAGCATGAGCTCTACCACCGCTTGAACCGATGAT	185
QY	56	SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIleGln	75
DB	186	GGTGACGAGACGATGGGCTTCCATTCACCTCTGGACGATGATGTCATACAG	245
QY	76	SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle	95
DB	246	TCUTTA---CCAAAGGCTTATGAGGAAATCGTACTGGAGCGTCTGAGAACTGATC	302
QY	96	GlyGluValLysAspIleMetPheAsnPheLysSerLysGluAspGlyGly-----	112
DB	303	GGGGAAGTAAAGAAC--ATGTTCAATTCGATGTCATTAGAAGATCGAGAGTTAATGAGT	359
QY	113	-----AsnAspLeuGlnArgLeuLeuValAspValGluArgLeuGlyIle	130
DB	360	CCGCTCAATGATCTCATTCACGCGCTTGGATTGTCGACAGCGCTTGAACGTTTGGGATC	419
QY	131	AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn	150
DB	420	CATGACATTTCAAGATGAGATAAATCGCGCTTGATATGTTTACAGTTATGGGCG	479
QY	151	GluLysGlyIleGlyCysGlyArgGluSerValThrAspLeuAsnSerThrAlaLeu	170
DB	480	GAATAATGGCATCGATCGGGAGGAGAGTGTGTTACTGATCTGAACCTCAACTCGGTTG	539
QY	171	GlyLeuArgThrIleArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe	190
DB	540	GGGCTTCGAACCTCAGACTACACGATACCCGGTGTCTCAGATGTTTTCAAGCTTTTC	599
QY	191	LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyIleArg	210
DB	600	AAAGGCCAAATGGGCGAGTTTCTCTCTGAAATATTTACAGACATGAAGAGATCAGA	659
QY	211	GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp	230
DB	660	GGCGTTCTGAATTTATTCGGGCGCTCCCTCAITTCCTTCCAGGGGAGAAAAATTATGAT	719
QY	231	GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer	250
DB	720	GAGGCTGAAATCTTCTCTACCAAAATATTTAAAGAGCGCTTCGAAAAGATTCCGCTCTC	779

QY	251	SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTrpHisThrAsnLeuPro	270
DB	780	AGT---CTTTCGCGAGAGATCGGGACGTTTGGATATGTTGGCACACATATTGCCG	836
QY	271	ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsnLysAsnAla	290
DB	837	CGATTGGGAAGCAAGGAATTACATCCAAGTCTTTGGACAGGACACTGAGAACAACGAAGTCA	896
QY	291	-----AlaGluLysLeuLeuAlaLysLeuGluPheAsnIlePheHisSer	307
DB	897	TATGTCAAGAGCAAAAACCTTTAGAACTCGCAAAATTGGAGTTCAACATCTTTCAATCC	956
QY	308	LeuGlnGluArgGluLysHisValSerArgTyrTrpLysAspSerGlySerProGlu	327
DB	957	TTACAAAGAGAGGAGTTAGAAAGTCTGGTCAGATGGTGAAGAAGATCGGGTTTCTGAG	1016
QY	328	MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe	347
DB	1017	ATGACCTTCTGCCGACATCGTCACGTGGAATACTACACTTTGGCTTCTGCAATGGCTTC	1076
QY	348	GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal	367
DB	1077	GAGCCTCAACATCTCTGGATTCAGACTCGGCTTTGCCAAGACCGTGTCTATATCAGGTT	1136
QY	368	LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr	387
DB	1137	CTTGACGATATGTACACACCTTCGCCACAGTAGACGAGCTGGAACCTCTTCACAGCGACA	1196
QY	388	IleLysArgTrpAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr	407
DB	1197	ATGAAGAGATGGATCGCTCCGTCTCGATAGATTCCTTCCAGAATATATGAAGAGGTGATC	1256
QY	408	MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg	427
DB	1257	ATAGCGGTTTACGACACCGTAATGAATGGTTCGAGAGCAGAGAGGCTCAAGGCCGA	1316
QY	428	AspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGlu	447
DB	1317	GATACGCTCACATATGCTCGGGAAGCTTGGAGGCTTATATTGATTCGTATATGCAAGAA	1376
QY	448	AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys	467
DB	1377	GCAAGGTGATCGCCACTGTTTACCTGCCCTCTCTTGTAGTACTACGAGAAATGGGAAA	1436
QY	468	ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe	487
DB	1437	GTTAGCTGTGTCATCGCATATCCGCATTTGCAACCCATTTCTGACATGGACATCCCTTT	1496
QY	488	ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle	507
DB	1497	CCTGATCATATCTCAAGGAAGTTGACTTCCCATCAAGCTTAACGACTTGGCATGTGC	1556
QY	508	IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu	527
DB	1557	ATCCTTCGATTACAGGTTGATACGGGTGCTCAAGGGCGGACAGGGCTCGTGAGAGAA	1616
QY	528	AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu	547
DB	1617	GTTTCTCTATATCATGTTATGAAGACATCTCTGGAGTATCAGAGGAAGATGCTCTC	1676
QY	548	AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLys	567
DB	1677	GATCATATCAACGCCATGATGATGACGTAAATCAAAGGATTAATTTGGGAATCTCTCAA	1736
QY	568	ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTrp	587
DB	1737	CCAGACATCAATGTTTCCCATCTCCGCGAAGAAACATGCTTTTGACATCCCGAGAGCTTC	1796
QY	588	HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu	607
DB	1797	CATTAGCGCTACAAATACCGAGAGCGCTACAGGTTGCCACGTTGAAACGAAGAGTTTG	1856

QY 608 ValMetArgThrValIleGluProValProLeu 618  
Db 1857 GTCACGAGAACCTCTGTAATCTGTGCTTGT 1889

RESULT 5  
AAAF73412  
ID AAF73412 standard; cDNA; 2186 BP.  
AC AAF73412;  
XX  
XX 30-APR-2001 (first entry)  
XX  
XX Grand fir monoterpene synthase coding sequence fragment SEQ ID NO: 66.  
DE  
XX Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
KW terpinolene synthase; insect resistance; nutrition; ss.  
XX  
XX Abies grandis.  
OS  
XX WO200107565-A2.  
PN  
XX 01-FEB-2001.  
PD  
XX 24-JUL-2000; 2000WO-US020264.  
PF  
XX 26-JUL-1999; 99US-00360545.  
PR  
XX (UNIW) UNIV WASHINGTON STATE RES FOUND.  
PA  
XX  
XX Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
PI  
XX WPI; 2001-182782/18.  
DR  
XX P-PSDB; AAB69391.  
XX  
XX New nucleic acid encoding monoterpene synthases, for increasing terpene  
PT synthesis in plants, e.g. for increasing resistance to pests or for  
PT treatment of cancer.  
XX  
XX Claim 13; Page 151-154; 175pp; English.  
XX  
XX The present invention provides the protein and coding sequences of  
CC monoterpene synthases from the grand fir. These include (-)-camphene  
CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
CC limonene, (-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
CC and pinene synthase. The sequences can be used to produce transgenic  
CC plants expressing high levels of the enzymes, resulting in levels which  
CC are useful in protecting against and treating cancers, and to confer  
CC insect resistance on plants  
XX  
XX Sequence 2186 BP; 722 A; 440 C; 472 G; 552 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 5,24e-224 Length: 2186  
Score: 2264.00 Matches: 435  
Percent Similarity: 80.53% Conservative: 78  
Best Local Similarity: 68.29% Mismatches: 98  
Query Match: 69.64% Indels: 26  
DB: 4 Gaps: 9

US-10-025-145A-65 (1-618) x AAF73412 (1-2186)

QY 1 MetAlaLeuSerIleThrProLeuValSerArgSerCysLeu----- 15  
Db 34 ATGGCTCTGGTTCTTCCGACCC-----AAATCTGCCTGCACAAATCGTTGATC 84

QY 16 ---SerSerSerHisGluIleIleAlaLeuArgThrIleProThrLeuGlyIleCys 34  
Db 85 AGGCTCTACTCATATGAGCTCAAGCCTCTGGCGAGAACCATCCCAACTTTGGAATGT 144

QY 35 ArgProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThr 54  
Db 145 AGGCGAGGGAATCTTTCACACCTTCTGTGAGCATGAGTTGACACCGCTGTATCTGAT 204

QY 55 AspSerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTyrPAspAspPheIle 74  
Db 205 GATGGTCTCAAAAGACGCTAGGTGACTATCATTTCCAATCTCTGGGACGACGATTCATA 264

QY 75 GlnSerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeu 94  
Db 265 CAGTCTCTA---TCAACGCCCTTATGGGAGCCTTCTTACCAGAACGCTGTGAGAAATG 321

QY 95 IleGlyGluValLysAspIleMetPheAsnPhelLysSerLeuGluAspGlyGly----- 112  
Db 322 ATTGGGAGAGTGAAGAG---ATGTTCAATTCATGCCATCGGAAGATGGAAATCAATG 378

QY 113 -----AsnAspLeuLeuGlnArgLeuLeuValAspAspValGluArgLeuGly 129  
Db 379 AGTCCCTCAATGATCTTATTGAACGACTTTGGATGGTCGATAGCGTTGAACGTTGGGG 438

QY 130 IleAspArgHisPheLysLysGluIleIleThrAlaLeuAspTyrValAsnSerTyrTrp 149  
Db 439 ATTGATAGACATTTCAAAAAAGAGATAAATCAGCCCTTGATTATGTTACAGTTATG 498

QY 150 AsnGluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAla 169  
Db 499 AACGAAAAAGGTATTTGGATCGGTAGAGATAGTGTCTTCTGATGTCACACTCGACTGCC 558

QY 170 LeuGlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerValLeuAsnVal 189  
Db 559 TCGGGGTTTCGAACCTCTTCGCCTACACGGATACAGTGTCTCTCAGAGGTTTGAAGA 618

QY 190 PheLysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGluIle 209  
Db 619 TTTCAAGACCAAAATGGGCAGTTTGCATTTCTCTCTAGTACAAA---GAGAGAGACATC 675

QY 210 ArgGlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMet 229  
Db 676 AGAACCGTTCTGAATTTATTCGGGCTTCTTCATTCCTCTGGGAGAAATGTTATG 735

QY 230 AspGluAlaGluThrPheSerThrLysTyrIleuArgGluAlaLeuGlnLysIleProAla 249  
Db 736 GAAGAGGCTGAAATTTCTCTCAAGATATTGAAAGACGCGTGAAGAAAGATTCGCGTC 795

QY 250 SerSerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTyrHisThrAsnLeu 269  
Db 796 TCCAGT---CTTTCACAGAAATAGACATACATTTGGAATATGTTGGCACACAAATATG 852

QY 270 ProArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThr----- 285  
Db 853 CCAAGATTGGAAACAAGGAATTACTTAGATGTATTGGACATCTCACCAGTCCATGGCTC 912

QY 286 LysAsnLysAsnAla-----AlaGluLysLeuLeuGluAlaLysLeuGlu 301  
Db 913 AAGAAGAAAAGGACGCAATATCTGGACAGCGAAAAGCTTTTAGAACTCGCAAAATGGAG 972

QY 302 PheAsnIlePheHisSerLeuGlnArgGluLeuLysHisValSerArgTrpTrpLys 321  
Db 973 TTCACATCTTTCATCTCCCTTCACAGAGAGAGTTACAGTATCTCTCCAGATGTGGATA 1032

QY 322 AspSerGlySerProGluMetThrPheCysArgHisArgHisValGluTyrTyrAlaLeu 341  
Db 1033 CATTCGGGTTTGCCTGAACTGACCTTTGGTCGGCATCGTCACGTGGGAATACTACACCTG 1092

QY 342 AlaSerCysIleAlaPheGluProGlnHisSerGlyPheArgLeuGlyPheThrLysMet 361  
Db 1093 AGCTCTTGCAATTGGCACTGAGCCCAACATTCGATTCAGATTGGGCTTTGCCAAAACG 1152

QY 362 SerHisLeuIleThrValLeuAspAspMetTyrAspValPheGlyThrValAspGluLeu 381  
Db 1153 TGTTCATCTTATCAGGGTCTTGGACGATATCTAGACACTTTCGGAAACGATGGATGAATC 1212

QY 382 GluLeuPheThrAlaThrIleLysArgTrpAspProSerAlaMetGluCysLeuProGlu 401  
Db 1213 GAATCTCTCAACAGGCGAGTTAGAGATGGAATCCGTCGGGAGAAAGACCGCTCCCGAA 1272

QY 402 TyrMetLysGlyValTyrMetMetValTyrHisThrValAsnGluMetAlaArgValAla 421  
 Db 1273 TATATGAAGAATCTACATGGCACTCTACGAGCCTTAACGTACATGGCGGAGGCA 1332  
 QY 422 GluLysAlaGlnGlyArgAspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPhe 441  
 Db 1333 GAGAAGACACAAGCGGACACAGCTCAATTATGTAGAAAGCTTGGGAAGTTTATCTT 1392  
 QY 442 AspSerTyrMetGlnGluAlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGlu 461  
 Db 1393 GATTTCGTATACCAAGAAAGTGGATGGCGAGCGTTATCTGCAACTTTCGAGGAG 1452  
 QY 462 TyrLeuGluAsnGlyLysValSerSerAlaHisArgProCysAlaLeuGlnProIleLeu 481  
 Db 1453 TACTTAGAGAACCGAGAGTTAGCTCTGCTCATCTGTCAGCGCATTTGACACCCCTCTG 1512  
 QY 482 ThrLeuAspIleProPheProAspHisIleLeuLysGluValAspPheProSerLysLeu 501  
 Db 1513 ACATTGGACGTACCGCTCTCTGATGACGCTTGAAGGGAATAGATTTTCCATCGAGATT 1572  
 QY 502 AsnAspLeuIleCysIleIleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAsp 521  
 Db 1573 AATGATTTGGCATCT 1632  
 QY 522 ArgAlaArgGlyGluAlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeu 541  
 Db 1633 AGGACCGAG 1692  
 QY 542 ThrGluGluAspAlaLeuAsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeu 561  
 Db 1693 ACAGAGGAAGATGCTCTCAATCATATCAATGCTCATGATCAACGACACATAATCAAGAATTA 1752  
 QY 562 AsnTrpGluLeuLeuLysProAspAsnSerValProIleThrSerLysLysHisAlaPhe 581  
 Db 1753 AATTGGGAATCTCTCAACCCCATAGCAATATCCCAATGCTGACGAGGAACATGCTTAT 1812  
 QY 582 AspIleSerArgValTrpHisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsn 601  
 Db 1813 GAGATAACAGAGAGCTTCCACCAACTTTACAAATATAGAGATGGCTTCAGCGTTGCCACT 1872  
 QY 602 ValGluThrLysSerLeuValMetArgThrValIleGluProValProLeu 618  
 Db 1873 CAAGAAACGAAAGTTTGGTGAGGAGACGGCTCTTGAACCACTGCTGCTCTCT 1923  
 RESULT 6  
 ID AAX08643 standard; cDNA; 2196 BP.  
 AC AAX08643;  
 DT 27-SEP-1999 (first entry)  
 XX Myrcene synthase gene.  
 DE Myrcene synthase gene.  
 XX Myrcene synthase; limonene synthase; pinene synthase; flavour; aroma;  
 KW defense; plant seed; oil; meal; ss.  
 XX Abies grandis.  
 XX Key Location/Qualifiers  
 CDS 69..1952  
 FT /\*tag= a  
 FT /product= "Myrcene synthase"  
 XX WO9902030-A1.  
 XX 21-JAN-1999.  
 XX 10-JUL-1998; 98WO-US014528.  
 XX 11-JUL-1997; 97US-0052249P.  
 XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX Bohlmann J, Steele CL, Croteau RB;  
 PI WPI; 1999-120396/10.  
 DR P-PSDB; AAW85700.  
 XX New isolated gymnosperm monoterpene synthase DNA - obtained from Grand  
 fir (Abies grandis), used to provide plants with modified production of  
 monoterpenes, e.g. myrcene, limonene or pinene.  
 PT Claim 9; Page 69-72; 121pp; English.  
 XX Nucleotide sequences encoding myrcene synthase, limonene synthase and  
 pinene synthase from Grand fir may be incorporated into any organism  
 (e.g. intact plant, animal, microbe), or derived cell culture that  
 produces geranyl diphosphate for the production of the aforementioned  
 enzymes or their products. The sequences when expressed in transfected  
 cells may also be used for the production or modification of flavour and  
 aroma properties, improvement of defense capability, and the alteration  
 of other ecological interactions mediated by myrcene, limonene, pinene,  
 or their derivatives. In particular they can be used for the production  
 of plant seeds for the extraction of oil or meal  
 XX Sequence 2196 BP; 684 A; 416 C; 468 G; 628 T; 0 U; 0 Other;  
 SQ Alignment Scores:  
 Pred. No.: 2,01e-217 Length: 2196  
 Score: 2200.50 Matches: 427  
 Percent Similarity: 80.38% Conservative: 77  
 Best Local Similarity: 68.10% Mismatches: 106  
 Query Match: 67.69% Indels: 17  
 DB: Gaps: 7  
 US-10-025-145A-65 (1-618) x AAX08643 (1-2196)  
 QY 1 MetaAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15  
 Db 69 ATGGCTCTGTTTCTTATCTACCGTTGGTCTCGAATCTTGGCTGGCAAGTGTGTATC 128  
 QY 16 SerSerSerHisGluIleLysAlaLeuArgArgThrIleProThrLeuGlyIleCysArg 35  
 Db 129 AGTTCAATTCAATGAACATAAAGCTCCCTATAGAACAAATCCCAATCTTGAATGCGTAG 188  
 QY 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55  
 Db 189 CGAGGGAATCTGTACGCTTCCATGACATCAGTTTGGCCACCGCTGCACCTGATGAT 248  
 QY 56 SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIleGln 75  
 Db 249 GGTGTACAAAGACGCGATAGGTACTACCATTCATATCTGGGACCGATGATTTTCATACAG 308  
 QY 76 SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95  
 Db 309 TCTCTA---TCACGCTTATGGGAACCTCTTACCAGAACGCTGCTGAGAGATTAAT 365  
 QY 96 GlyGluValLysAspIleMetPheAsnPheLysSerLeuGluAspGlyGly----- 112  
 Db 366 GTGAGGTAAGAAGATA---TTCAATTCAATGTACTGTGATGATGGAAGATTATATGAGT 422  
 QY 113 -----AsnAspLeuLeuGlnArgLeuLeuValAspValGluArgLeuGlyIle 130  
 Db 423 TCTTTAATGATCTCATGACAGCGCTTTGGATAGTGTAGCGTTGAACGTTTGGGGATA 482  
 QY 131 AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150  
 Db 483 GCTAGACATTTCAAGACGAGATACATCAGCTCTGGATTATGTTTCCGTTACTGGGAG 542  
 QY 151 GluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170  
 Db 543 GAAACGCGCATTCGATGTGGGAGAGACAGTATTGTTACTGATCTCAACTCAACTGCGTTG 602  
 QY 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190

603	GGGTTTCGAACCTCTTCGATTACACGGGTACACTGTATCTCCAGAGGTTTAAAAAGCTTTT	662
191	LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGluIleArg	210
663	CAAGATCAAAATGACACGATTGTATGCTCCCCCGT---CAGACAGAGGTCAGATCAGA	719
211	GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp	230
720	AGCGTCTTAACTTATATCGGGCTTCCTCATTCCTCCCTCCCTGGTGAGAAAGTTATGGAA	779
231	GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer	250
780	GAAGCTGAAATCTTCTCCACAGATATTGGAAGAAGCTCTACAAAAGATTCAGTCTCC	839
251	SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTyrHisThrAsnLeuPro	270
840	GCT---CTTTCAAGAGATAAAGTTGTATGGAATATGGCTGGCACACAAATTTGCCA	896
271	ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLys-----Asn	287
897	AGATTGGAAGCAGAAATTACATAGACACACTTCGAGAAAGACACACAGTCGATGGCTCAAT	956
288	LysAsnAlaAlaGluLysLeuLeuGluLeuAlaLysLeuGluPheAsnIlePheHisSer	307
957	AAAAATGCTGGGAAGAAGCTTTTAGAACTTCGAAATTTGGAGTTCAATATATTTAACTCC	1016
308	LeuGlnGluArgGluLeuLysHisValSerArgTyrTrpLysAspSerGlySerProGlu	327
1017	TTACACAAAGAAGAAATTACAATATCTTTGAGATGGTGGAAAGAGTCGATTTCGCTTAAA	1076
328	MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe	347
1077	TTGACATTTCGTCGCATCGTCATGCGAATTTCTACACTTTGGCGCTCTTGTTATGGCATT	1136
348	GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal	367
1137	GACCCAAAACATCTCGCATTCAGACTAGGCTTCGCCAAAATGTGTCATCTTGTCACAGTT	1196
368	LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr	387
1197	TTGACCAATATTACGACACTTTTGGAAAGATTCAGAGCTTGAACTCTTCACATCTGCA	1256
388	IleLysArgTrpAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr	407
1257	ATTAAGAGATGGAATTTCATCAGATAGAACACCTTCACAGATATATGAAATGTGTAC	1316
408	MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg	427
1317	ATGCTCGTGTGTGAACTGTAAATGAACTGACACGAGGCGGAGAGACTCAAGGGAGA	1376
428	AspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGlu	447
1377	AACACTCTCAACTATGTTGNAAGGCTTGGAGGCTTATTTTGATTCAATATATGGAAGA	1436
448	AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys	467
1437	GCAAAATGGATCTCTAATGGTTATCTGCAATGTTTGAAGAGTACCATGAGAATGGGAAA	1496
468	ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe	487
1497	GTGAGCTCTGCATATCGGTAGCAACATTCGAACCCATCTCTCACTTTGAATGCATGGCTT	1556
488	ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle	507
1557	CCTGATTACATCTTGAAAGGAATTTGATTTTCCATCCAGGTTTCAATGATTTGGCATCGTCC	1616
508	IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu	527
1617	TTCTCTCGGCTACGAGGTGACACAGCTGCTACAAGCCGATAGGATCGTGGTGAAGA	1676
528	AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluAspAlaLeu	547
1677	GCTTCGTGTATCATGTATTATGAAGACAACTCTGGATCAACCGAAGAGATGCCTC	1736

Qy	548	AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLeuLys	567
Db	1737	AATCATATCATGCGATGGTCAATGACATAATCAAGAAATTAATTTGGGAACCTTCTAAGA	1796
Qy	568	ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTrp	587
Db	1797	TCCAAACACCAATATCCAAATGCTGGCCCAAGAAACATGCTTTTGACATAACAAGAGCTCTC	1856
Qy	588	HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu	607
Db	1857	CACCATCTACATATATCGAGATGGCTTAGTGTGCCCAACAAGGAAACAAAAAATTG	1916
Qy	608	ValMetArgThrValIleGlu	614
Db	1917	GTTATGGAACACTTCCTTGAA	1937
RESULT	7		
AAA38927			
ID	AAA38927	standard; DNA; 2196 BP.	
XX			
AC	AAA38927;		
XX			
DT	25-AUG-2000	(first entry)	
XX			
DE	Grand fir myrcene synthase DNA sequence SEQ ID NO:29.		
XX			
KW	Synthase; protein co-ordinate data; active site; modification; terpenoid;		
KW	3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;		
KW	isoprenoid; breeding programme; fragrance; flavour; pheromone;		
KW	defensive agent; pigment; antitumour; steroid hormone;		
KW	signal transduction pathway; bile acid; affinity purification;		
KW	photoreceptor; enzymatic synthesis; nutrient supplement;		
KW	immunological reagent; ds.		
XX			
OS	Abies grandis.		
XX			
PN	WO200017327-A2.		
XX			
PD	30-MAR-2000.		
XX			
PF	17-SEP-1999; 99WO-US021419.		
XX			
PR	18-SEP-1998; 98US-0100993P.		
PR	22-APR-1999; 99US-0130628P.		
PR	23-AUG-1999; 99US-0150262P.		
XX			
PA	(KENT ) UNIV KENTUCKY RES DEPT.		
PA	(SALK ) SALK INST BIOLOGICAL STUDIES.		
XX			
PI	Chappell J, Manna KR, Noel JP, Starks CM;		
XX			
DR	WPI; 2000-292839/25.		
DR	P-PSDB; AAY90842.		
XX			
PT	Novel terpene synthase enzymes, useful for producing terpene		
PT	hydrocarbons, e.g. fragrances or antitumor agents, are derived from known		
PT	enzymes by specific amino acid alterations.		
XX			
FS	Disclosure; Page 387-390; 450pp; English.		
XX			
CC	The present invention describes an isolated terpene synthase (I)		
CC	comprising a region with at least 20% identity to region 265-335 of a 548		
CC	amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha		
CC	-carbon atoms (alphaC) that have interatomic distances, between each		
CC	other, within tabulated ranges, have a centre point (within a sphere of		
CC	radius 2.3 Angstrom ) within tabulated ranges, and have an ordered		
CC	arrangement of R groups (defining a side chains), excluding specific		
CC	tabulated arrangements (tables given in the specification). (I), and		
CC	related enzymes, are used to produce a wide range of terpenoids (e.g.		
CC	cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,		
CC	flavours, pheromones, defensive agents, pigments, antitumour agents,		
CC	components of signal transduction pathways, precursors of steroid		

CC hormones and bile acids, as photoreceptors and as co-factor side chains.  
 CC Some synthases with little or no catalytic activity (and nucleic acids  
 CC encoding them) are used as controls in the analysis of products formed by  
 CC enzymatic synthesis; as nutrient supplements; for affinity purification  
 CC of isoprenoids; or to develop immunological reagents or nucleic acids for  
 CC monitoring expression of terpene synthase or inheritance of the gene in  
 CC plant breeding programs. The new synthases may produce novel terpene  
 CC products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 2196 BP; 684 A; 416 C; 468 G; 628 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2,01e-217 Length: 2196  
 Score: 2200.50 Matches: 427  
 Percent Similarity: 80.38% Conservative: 77  
 Best Local Similarity: 68.10% Mismatches: 106  
 Query Match: 67.69% Indels: 17  
 DB: 3 Gaps: 7

US-10-025-145A-65 (1-618) x AAA38927 (1-2196)

QY 1 MetAlaLeuSerIleThrProLeuValSerArgSerCysLeu----- 15  
 Db ATGGCTCTGTTCTATCTCACCGTTGGCTTGGAAATCTGCCTGGCGCAAGTCGTTGATC 128  
 QY 16 SerSerSerHisGluIleValAlaLeuArgArgThrIleProThrLeuGlyIleCysArg 35  
 Db 129 AGTTCGAATTCATGAACATGAAGCCCTCCCTATAGAACCAATCCCAATCTTGAATGGCTAGG 188  
 QY 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55  
 Db 189 CGAGGGAATCTGTACGCTTCCATGAGCATCAGTTTGGCCACCGCTGCACCTGATGAT 248  
 QY 56 SerValGlnArgValGlyAsnTyrHisSerAsnLeuThrAspAspPheIleGln 75  
 Db 249 GGTGTACAAAGACGATAGTGTACTACCATTCATATCTGGACGATGATTTTCATACAG 308  
 QY 76 SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95  
 Db 309 TCTCTA---TCAACGCTTATGGGGAACCTCTTACAGGAAGCTGTGAGATTAAT 365  
 QY 96 GlyGluValLysAspIleMetPheAsnPheLysSerLeuGluAspGlyGly----- 112  
 Db 366 GTGGAGTAAAGAGATA---TTCATTTCAATCTGATGATGATGATGAAGATTAAAGAT 422  
 QY 113 -----AsnAspLeuGlnArgLeuLeuValAspValGluArgLeuGlyIle 130  
 Db 423 TCCTTTAATGATCTCATGCAACGCTTTGGATAGTCGATAGCGTTGAACGTTTGGGGATA 482  
 QY 131 AspArgHisPheIleLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150  
 Db 483 GCTAGACATTTCAAGAACGAGATAACATCAGCTCTGGATTATGTTTCCGTTACTGGGAG 542  
 QY 151 GluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170  
 Db 543 GAAACGCGATTGGATGTGGAGAGACAGATTGTTTACTGATCTCAACTCAACTGCGTTG 602  
 QY 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190  
 Db 603 GGGTTTCGAACCTCTTCGATTTACACGGGTACCTGTATCTCCAGAGTTTTAAAGCTTTT 662  
 QY 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGluIleArg 210  
 Db 663 CAGATCAAAATGGACAGATTGTATGCTCCCCGGT---CAGACAGAGGTGAGATCAGA 719  
 QY 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp 230  
 Db 720 AGCGTTCTTAACATTATATCGGGCTTCCCTCATTCGCTTCCCTGCTGAGAAAGTTATGGA 779  
 QY 231 GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer 250  
 Db 780 GAAGCTGAAATCTCTCCACAGATATTTGAAAGAGAGCTCTACAAAAGATTCACAGTCTCC 839

QY 251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTyrHisThrAsnLeuPro 270  
 Db 840 GCT---CTTTCAAGAGATAAAGTTTGTATGGAATATGGCTGGCACAAATTTGCCA 896  
 QY 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLys-----Asn 287  
 Db 897 AGATTGGAAGCAAGAAATTTACATAGACACATTTGAGAAGACACACAGTCATGGCTCAAT 956  
 QY 288 LysAsnAlaAlaGluLysLeuLeuAlaLysLeuGluPheAsnIlePheHisSer 307  
 Db 957 AAAAATGCTGGGAAGAGCTTTTAGAACTTGCAAAATTTGAGTTCATATATTAACTCC 1016  
 QY 308 LeuGlnGluArgGluLeuLysHisValSerArgTyrTyrLysAspSerGlySerProGlu 327  
 Db 1017 TTACACAAAGGAATTCATATCTTTGAGAGTGGTGAAGAGTGGGATTTGCCCTAAA 1076  
 QY 328 MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347  
 Db 1077 TTGACATTTGCTCGGCATCGTCATGTGGAATTCACACTTTGGCCTCTTTGATTGCCATT 1136  
 QY 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal 367  
 Db 1137 GACCCAAAACATTTCTGCAATTCAGACTAGCTTTCGCCAAAATGTGTCTCTTGCAGATT 1196  
 QY 368 LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387  
 Db 1197 TTGGACGATATTACGACACTTTTGGACGATTGACGACTTGAACCTTTCACATCGCA 1256  
 QY 388 IleLysArgTyrAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407  
 Db 1257 ATTAAGAGATGCAATTCATCAGATAGAACACCTTCCAGAAATATATGAAATGTGTATC 1316  
 QY 408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427  
 Db 1317 ATGGTCTGTTTGAACACTGTAAATGAACACTGACACGAGCGCGGAGAAAGACTCAAGGAGA 1376  
 QY 428 AspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGlu 447  
 Db 1377 AACACTCTCACTATGTTTCGAAAGGCTTGGAGGCTTATTTGATTCATATATGGAAGAA 1436  
 QY 448 AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys 467  
 Db 1437 GCMAATGATCTCTAATGTTATCTGCCAATGTTTGAAGAGTAGTACCATGAGAATGGAAA 1496  
 QY 468 ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe 487  
 Db 1497 GTGAGCTCTGCATATCGCGTAGCAACATTGCAACCCATCTCTCACTTTGAATGATGGCTT 1556  
 QY 488 ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle 507  
 Db 1557 CTTGATTACATCTTGAAGGGAATGATTTTCCATCAGGTTCAATGATTTGGCATCGTCC 1616  
 QY 508 IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu 527  
 Db 1617 TTCCTTCGGCTACGAGGTGACACACGCTGCTCAAGCGCGAGTAGGATCGTGGTGAAGAA 1676  
 QY 528 AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu 547  
 Db 1677 GCTTCGTGATATCATGTTTATGAAAGACAATCTCTGGATCAACCGAAGAGATGCCCTC 1736  
 QY 548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLeuLys 567  
 Db 1737 AATCATATCAATGCCATGCTCAATGACATAATCAAGAATTAATTTGGGAATCTTCAAGA 1796  
 QY 568 ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTrp 587  
 Db 1797 TCACACGACATATTCATATGCTGCCAAGAACATGCTTTTGCATACACAGAGCTCTC 1856  
 QY 588 HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu 607  
 Db 1857 CACCATCTCTACATATATCGATGCTTTAGTTGTTGCCCAACAGGAACCAAAAAATTG 1916

Qy 608 ValMetArgThrValIleGlu 614  
 Db 1917 GTTATGGAACACACTCCTTGAA 1937

RESULT 8  
 AAF73371  
 ID AAF73371 standard; cDNA; 2196 BP.  
 AC AAF73371;  
 XX  
 XX 30-APR-2001 (first entry)  
 DT  
 XX  
 DE Grand fir myrcene synthase coding sequence SEQ ID NO: 1.  
 XX  
 XX Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
 KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
 KW terpinolene synthase; insect resistance; nutrition; ss.  
 XX  
 OS Abies grandis.  
 XX  
 XX WO200107565-A2.  
 FN  
 XX  
 XX 01-FEB-2001.  
 PD  
 XX  
 XX 24-JUL-2000; 2000WO-US020264.  
 PF  
 XX  
 XX 26-JUL-1999; 99US-00360545.  
 PR  
 XX  
 XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 PA  
 XX  
 XX Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
 PI  
 XX WPI; 2001-182782/18.  
 DR  
 XX P-PSDB; AAB69370.  
 XX  
 PT New nucleic acid encoding monoterpene synthases, for increasing terpene  
 PT synthesis in plants, e.g. for increasing resistance to pests or for  
 PT treatment of cancer.  
 XX  
 XX Claim 28; Page 103-106; 175pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of  
 CC monoterpene synthases from the grand fir. These include (-)-camphene  
 CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 CC limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 CC and pinene synthase. The sequences can be used to produce transgenic  
 CC plants expressing high levels of the enzymes, resulting in levels which  
 CC are useful in protecting against and treating cancers, and to confer  
 CC insect resistance on plants  
 XX  
 XX Sequence 2196 BP; 684 A; 416 C; 468 G; 628 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2,01e-217 Length: 2196  
 Score: 2200.50 Matches: 427  
 Percent Similarity: 80.38% Conservative: 77  
 Best Local Similarity: 68.10% Mismatches: 106  
 Query Match: 67.69% Indels: 17  
 DB: 4 Gaps: 7

US-10-025-145A-65 (1-618) x AAF73371 (1-2196)

Qy 1 MetAlaLeuSerIleThrProLeuValSerArgSerCysLeu----- 15  
 Db 69 ATGGCTCTGGTTCTATCTACCGTTGGCTTCGAAATCTTGCCTGCGCAAGTCGTTGATC 128

Qy 16 SerSerSerHisGluIleLysAlaLeuArgThrIleProThrLeuGlyIleCysArg 35  
 Db 129 AGTTCAATTCATGACATAAGACCTCCCTATGAAACAATCCCAAAATCTTTGGAATCGGTAGG 188

Qy 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55  
 Db 189 CGAGGGAATCTGTACCCCTTCCATGAGCATCAGTTTGGCCACCGCTGCACCTGATGAT 248

Qy 56 SerValGlnArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIleGln 75  
 Db 249 GGTGTACAAAGACGATAGGTGACTACCATTTCCAATATCTGGGACGATGATTTTCATACAG 308

Qy 76 SerLeuIleSerThrProTyrGlyValaProAspTyrArgGluArgAlaAspArgLeuIle 95  
 Db 309 TCTCTA--TCAACGCCTTATGGGAAACCCCTCTTACCAGGAACGCTGCTGAGAGATAAAT 365

Qy 96 GlyGluValLysAspIleMetPheAsnPhelysSerLeuGluAspGlyGly----- 112  
 Db 366 GTGGAGGTAAAGAAGATA--TTCAATTCAATGATACCTGGATGATGGAGATTAATGAGT 422

Qy 113 -----AsnAspLeuLeuGlnArgLeuLeuValaAspValGluArgLeuGlyIle 130  
 Db 423 TCCTTTAAATGATCTCATGCAACGCTTTGGATAGTCGATAGCGTTGAACGTTGGGGATA 482

Qy 131 AspArgHisPhelysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150  
 Db 483 GCTAGACATTTCAAGAACGAGATAACATCAGCTCTGGATTATGTTTCCGTTACTGGGAG 542

Qy 151 GluLysGlyIleGlyCysGlyValGluSerValValThrAspLeuAsnSerThrAlaLeu 170  
 Db 543 GAAACCGCATTTGGATGTGGGAGAGACAGATTGTTACTGATCTCAACTCACTGCGTTG 602

Qy 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190  
 Db 603 GGGTTTCGAACCTCTTCGATTACACGGGTACACTGTATCTCCAGAGGTTTAAAGAGCTTTT 662

Qy 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyLeuArg 210  
 Db 663 CAAGATCAAAATGGACAGATTGTTATGCTCCCCCGGT---CAGACAGAGGGTGAGATCAGA 719

Qy 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp 230  
 Db 720 AGCGTTCTTAACATTATATCGGGCTTCCCTCATTCCTTCCCTGGTGAGAAAGTTATGGAA 779

Qy 231 GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer 250  
 Db 780 GAAGCTGAAATCTTCTCCACAAGATATTTGAAAGAAGCTCTACAAAAGATTCAGATCTCC 839

Qy 251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTrpHisThrAsnLeuPro 270  
 Db 840 GCT---CTTTCACAGAGATAAAGTTGTTATGGAATATGGCTGGCACACAAATTTGCCA 896

Qy 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLys-----Asn 287  
 Db 897 AGATTGGAAGCAAGAAATTACATAGACACACTTGCAGAAAGACACACAGTGCATCAAT 956

Qy 288 LysAsnAlaAlaGluLysLeuLeuGluAlaLysLeuGluPheAsnIlePheHisSer 307  
 Db 957 AAAAATGCTGGAAAGAGCTTTTAGAACTTGCAAAATTTGGAGTTCAATATATTTAACTCC 1016

Qy 308 LeuGlnGluArgGluLeuLysHisValSerArgTrpTrpLysAspSerGlySerProGlu 327  
 Db 1017 TTACACAAAGAGAAATTACAAATATCTTTTGAGATGGTGGAAGAGTCGGATTGCGCTAAA 1076

Qy 328 MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347  
 Db 1077 TTGACATTTGCTCGCATCGTCATGTCGAATTTACACTTTGGCCTCTTGTATTGCCATT 1136

Qy 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal 367  
 Db 1137 GACCAAAACATTTCTGCATTCAGACTAGGCTTCGCCAAAATGTGTCTCATCTTGTCCACAGTT 1196

Qy 368 LeuAspAspMetTyrAspValPheGlyThrValaAspGluLeuGluLeuPheThrAlaThr 387  
 Db 1197 TTGACCATATTTACGACACTTTTGGAACGATTCAGCGCTTGAACTCTTCCATCTGCA 1256

Qy 388 IleLysArgTrpAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407  
 Db 1257 ATTAAGAGATGGAATTCATCAGAGATAGAACACCTTCCAGATATATCAATATGTGTATC 1316



```
Db 591 GGGTTTCGAACTCTTCGATTACACGGGTACACTGTATCTCAGAGGTTTAAAGACTTTT 650
Qy 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGluLeuLeuLeu 210
Db 651 CAAGATCAAAATGACAGCTTTGTATGCTCCCGCGT---CAGACAGAGGGTGAGATCAGA 707
Qy 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLeuValMetAsp 230
Db 708 AGCGTTCCTTAACTTATATGCGGCTCCCTCATTTGCTTCCCTGGTGAGAAAAGTTATGGAA 767
Qy 231 GluAlaGluThrPheSerThrLysThrLysArgGluAlaLeuGlnLysIleProAlaSer 250
Db 768 GAAGCTGAAATCTCTCCACAGATATTGGAAGAGCTCTACAAAAGATTCCAGTCTCC 827
Qy 251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyThrHisThrAsnLeuPro 270
Db 828 GCT---CTTTCACAAGATTAAGTTTGTATGGNATATGCTGGCACACAAATTTGGCA 884
Qy 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLys-----Asn 287
Db 885 AGATTGGAAGCAAGAAATTACATAGACACACTTGAGAAAAGACACCAGTGCATGGCTCAAT 944
Qy 288 LysAsnAlaAlaGluLysLeuLeuAlaLysLeuGluPheAsnIlePheHisSer 307
Db 945 AAAAATGCTGGGAGAGAGCTTTTAGAAGCTTGCAAAATTTGGAGTTCAATATATTTATCC 1004
Qy 308 LeuGlnGluArgGluLysHisValSerArgTyrTrpLysAspSerGlySerProGlu 327
Db 1005 TTACACAAAAGGAATTACATATCTTTGAGATGGTGGAAAGAGTCGGATTTGCTTAAA 1064
Qy 328 MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347
Db 1065 TTGACATTTGCTGGCATCGTCATCGAATTTCTACACTTTGGCCCTCTTGTATTGCCATT 1124
Qy 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal 367
Db 1125 GACCCAAAACATCTCGCAATTCAGACTAGGCTTCGCCAAAATGTGTCACTTGTGCACAGTT 1184
Qy 368 LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluPheThrAlaThr 387
Db 1185 TTGGACGATATTACACACATTTTGGACGATTTGACGAGCTTGAACCTTTCACATCTGCA 1244
Qy 388 IleLysArgTyrAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407
Db 1245 ATTAAGAGATGGAATTCATCAGATAGAACACCTTCCAGAAATATATGAATGTGTGTAC 1304
Qy 408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427
Db 1305 ATGGTGGTGTGAAACTGTAAATGAAGTGCACGAGAGCGGAGAAAGACTCAAGGGAGA 1364
Qy 428 AspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGlu 447
Db 1365 AACACTCTCACTATGTTGGAAGGCTTGGGAGGCTTATTGATTTCATATATGGAAGAA 1424
Qy 448 AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys 467
Db 1425 GCAAAATGGATCTCTAATGGTTTATCTGCAACGTTTGAAGAGTACCATCAGAAATGGGAAA 1484
Qy 468 ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe 487
Db 1485 GTGAGCTTCGATATCGCGTAGCAATTCGAACCCATCCTCAGTTGAATCGATCGGCTT 1544
Qy 488 ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspIleCysIle 507
Db 1545 CTGTATTACATCTTGAAGGAATTGATTTTCCATCCAGTTCAATGATTGTCATCGTCC 1604
Qy 508 IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu 527
Db 1605 TTCTCTGGCTACGAGGTGACACACGCTGCTACAGGCCGATAGGATCGTGGTGAAGAA 1664
Qy 528 AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluAspAlaLeu 547
```

```
Db 1665 GCTTCGTGTATATCATGTTATATCAAGACAACTCCTGATCAACCGAAGAAAGTCCCTC 1724
Qy 548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLeuLys 567
Db 1725 AATCATATCAATGCCATGGTCAATGACATATCAAGAATTAATTTGGAACTTCTAAGA 1784
Qy 568 ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTrp 587
Db 1785 TCCAAACGACATATATCCAAATGCTGGCCAAGAAACATGCTTTTGACATAACAAGAGCTC 1844
Qy 588 HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu 607
Db 1845 CACCATCTACATATATCGAGATGGCTTTAGTTGTCACCAAGAAACAAAAAATTG 1904
Qy 608 ValMetArgThrValIleGlu 614
Db 1905 GTTATGGAACACTCTTGA 1925

RESULT 10
AAF73391
ID AAF73391 standard; cDNA; 2205 BP.
XX
AC AAF73391;
XX
DT 30-APR-2001 (first entry)
XX
DE Grand fir monoterpene synthase coding sequence fragment SEQ ID NO: 31.
KW Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;
KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;
KW terpinolene synthase; insect resistance; nutrition; ss.
XX
OS Abies grandis.
XX
PN WO200107565-A2.
XX
PD 01-FEB-2001.
XX
PF 24-JUL-2000; 2000WO-US020264.
PR 26-JUL-1999; 99US-00360545.
PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.
PI Steele CL, Bohlmann J, Croteau RB, Phillips MA;
XX
DR WPI: 2001-182782/18.
DR P-PSDB; AAB69380.
XX
PT New nucleic acid encoding monoterpene synthases, for increasing terpene
synthesis in plants, e.g. for increasing resistance to pests or for
treatment of cancer.
XX
PS Example 3; Page 135-137; 175pp; English.
XX
CC The present invention provides the protein and coding sequences of
monoterpene synthases from the grand fir. These include (-)-camphene
synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-
limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase
and pinene synthase. The sequences can be used to produce transgenic
plants expressing high levels of the enzymes, resulting in levels which
are useful in protecting against and treating cancers, and to confer
insect resistance on plants
XX
SQ Sequence 2205 BP; 703 A; 412 C; 463 G; 627 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.13e-217 Length: 2205
Score: 2197.50 Matches: 427
Percent Similarity: 80.38% Conservative: 77
Best Local Similarity: 68.10% Mismatches: 106
Query Match: 67.59% Indels: 17
Db: 4 Gaps: 7
```

US-10-025-145A-65 (1-618) x AAF73391 (1-2205)

Qy	1	MetalAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu-----	15
Db	57	ATGCTCTGGTTTCTATCTACCGTTGCTTCGAATCTTGCTCGCAAGTCGTTGATC	116
Qy	16	SerSerHisGluIleLysAlaLeuArgArgThrIleProThrLeuGlyIleCysArg	35
Db	117	AGTTCAATTCATGAACATAAGCCCTCTATAGACAATCCCAATCTTGGATGCGTAGG	176
Qy	36	ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp	55
Db	177	CGAGGAAATCTGTACACGCTTCCATGAGCATCAGTTTGGCCACCGCTGCACCTGATGAT	236
Qy	56	SerValGlnAspArgValGlyAsnTyrHisSerAsnLeuTyrAspAspPheIleGln	75
Db	237	GCTGTACAAAGACCATAGGTGTCACACCTTCCAAATATCTGGACGATGATTTCCATACAG	296
Qy	76	SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluAlaAspArgLeuIle	95
Db	297	TCTCTA---TCAAGCGCATATGGGAACCCCTCTTACCAGGAACGCTGAGAGATTAATT	353
Qy	96	GlyGluValLysAspIleMetPheAsnPhelysSerLeuGluAspGlyGly-----	112
Db	354	GTGGAGGTAAAGAAGATA---TCAATTCATATGATACCTGGATGTGGAGAGATTAAATGAGT	410
Qy	123	-----AsnAspLeuLeuGlnArgLeuLeuLeuValAspValGluArgLeuGlyIle	130
Db	411	TCCTTTTAATGATCTCATGCAACGCTTTGGATAGTCATGCGTTGAACGTTGGGATA	470
Qy	131	AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn	150
Db	471	GCTAGACATTTCAAGAACGAGATACACTCAGCTCGGATTATGTTTCCGTACTCGGAG	530
Qy	151	GluLysGlyIleGlyCysGluValGluSerValValThrAspLeuAsnSerThrAlaLeu	170
Db	531	GAATAACGGCATTTGATGTGGAGAGACAGATATTGTTACTGATCTCACTCAACTCGCGTTG	590
Qy	171	GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe	190
Db	591	GGGTTTCCAACTCTTCGATTACACGGGTACACTGATCTCAGAGAGTTTAAAAAGCTTTT	650
Qy	191	LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGluIleArg	210
Db	651	CAAGATCAAAATGGACAGTTTGTATGCTCCCGCGT---CAGACAGAGGCTGAGATCAGA	707
Qy	211	GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp	230
Db	708	AGCGTTCTTAACATTATATCGGGCTTCCTCATTTCCCTTCCCTGGTGAGAAAGTTATGGAA	767
Qy	231	GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer	250
Db	768	GAAGCTGAATCTCTCCACAGATATTTGAAAGAGCTCTACAAAAGATTCCAGTCTCC	827
Qy	251	SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTyrHisThrAsnLeuPro	270
Db	828	GCT---CTTTCAAGAGAGATAAAGTTTGTATGGAATATGCTGGCACACAAAATTGCCA	884
Qy	271	ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLys-----Asn	287
Db	885	AGATTGGACCAAGAAATTACATAGACACTTGAGAAAGACACCGATGCATGGCTCAAT	944
Qy	288	LysAsnAlaAlaGluLysLeuLeuGluLeuAlaLysLeuGluPheAsnIlePheHisSer	307
Db	945	AAAAATGCTGGGAAGAAGCTTTTAGAACTTGCAAAATGGAGTTTCAATATATATTTAATCC	1004
Qy	308	LeuGlnGluArgGluLeuLysHisValSerArgTrpTrpLysAspSerGlySerProGlu	327
Db	1005	TTTCAACAAAAGGAATTACAAATATCTTTTGAGATGGTGGAAAGAGTCGGATTTCGCTAAA	1064
Qy	328	MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe	347

Db	1065	TTGACATTGTGCTCGGCATCGTCATGTGGAAATCTACACTTTGGCGCTCTTGTTATGGCCATT	1124
Qy	348	GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal	367
Db	1125	GACCCAAAACATCTCGCAATTCAGACTAGGCTTCGCCAAATGTGTCACTTGTGCACAGTT	1184
Qy	368	LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr	387
Db	1185	TTGACGATATTTACGACACTTTTGGAAACGATTCACGAGCTTGAATCTTCACATCTGCA	1244
Qy	388	IleLysArgTrpAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr	407
Db	1245	ATTAAAGATGGAATTCATCAGATAGAACCCCTCCAGAATATATGAATGTGTGTAC	1304
Qy	408	MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg	427
Db	1305	ATGTCGTGTTGAACTGTAAATGAATGCACGAGAGCGGGGAGAAAGACTCAAGGAGA	1364
Qy	428	AspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGlu	447
Db	1365	AACACTCTCAACTATGTTCGAAAGGCTTGGGAGGCTTATTTTGATTCATATATGGAAGAA	1424
Qy	448	AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys	467
Db	1425	GCAAAATGGATCTCTAATGTTATCTGCCAACGTTTGAAGAGTACCATGAGAATGGGAAA	1484
Qy	468	ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe	487
Db	1485	GTGAGCTCTCGATATCGGTAGCAACATTCGACCCCACTCTCACTTTGAATGCATGGCGTT	1544
Qy	488	ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle	507
Db	1545	CCTGATTACATCTTGAAGGGAATGATTTTCCATCCAGGTTCAATGATTTGGCATCGTCC	1604
Qy	508	IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu	527
Db	1605	TTCTTCGGGTACGAGGTGACACACGCTGCTACAAGGCCGATAGGGATCGTGGTGAAGAA	1664
Qy	528	AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu	547
Db	1665	GCITTCGGTATATCATGTTATATGAAGAATCTCTGGATCAACCCGAAGAAGATGCCCTC	1724
Qy	548	AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLeuLys	567
Db	1725	AATCATATCAATGCCATGGTCAATGATCATATCAAGAATTAATTTGGAACTTCTAAGA	1784
Qy	568	ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTrp	587
Db	1785	TCCAAACGACAATATTCGAATGCTGGCCAGAAACATGCTTTTGACATAACAAGAGCTCTC	1844
Qy	588	HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu	607
Db	1845	CACATCTTACATATATCGAGATGGCTTTAGTGTGGCCAAACGAAGAAACAAAAAATTG	1904
Qy	608	ValMetArgThrValIleGlu	614
Db	1905	GTTATGAAACACTCTCTTGAA	1925
RESULT 11			
AAAF73421			
ID AAF73421 standard; cDNA; 1890 BP.			

RESULT 11

REF ID: AAF73421

ID AAF73421 standard; cDNA; 1890 BP.

XX

AC AAF73421;

XX  
XX

- 1 -

DT 30-APR-2001 (first entry)

XX

DE Grand fir monoterpene synt

XXI

KW Monoterpene synthase; gran

myrcene synthase; (-)-limo-

KW terpinolene synthase; insect  
VV

XX  
50  
Abies  
Picea  
Picea

Us. Abies grandis.

XX WO200107565-A2.  
 XX 01-FEB-2001.  
 XX 24-JUL-2000; 2000WO-US020264.  
 XX 26-JUL-1999; 99US-00360545.  
 XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 XX Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
 XX WPI; 2001-182782/18.  
 XX P-FSDB; AAB69393.  
 XX New nucleic acid encoding monoterpene synthases, for increasing terpene  
 XX synthesis in plants, e.g. for increasing resistance to pests or for  
 XX treatment of cancer.  
 XX Claim 18; Page 163-165; 175pp; English.  
 XX The present invention provides the protein and coding sequences of  
 XX monoterpene synthases from the grand fir. These include (-)-camphene  
 XX synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 XX limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 XX and pinene synthase. The sequences can be used to produce transgenic  
 XX plants expressing high levels of the enzymes, resulting in levels which  
 XX are useful in protecting against and treating cancers, and to confer  
 XX insect resistance on plants  
 XX  
 XX Sequence 1890 BP; 550 A; 381 C; 441 G; 518 T; 0 U; 0 Other;  
 XX  
 XX Alignment Scores:  
 XX Pred. No.: 6.11e-211 Length: 1890  
 XX Score: 2137.00 Matches: 427  
 XX Percent Similarity: 78.86% Conservative: 73  
 XX Best Local Similarity: 67.35% Mismatches: 114  
 XX Query Match: 65.73% Indels: 20  
 XX DB: 4 Gaps: 9  
 XX  
 XX US-10-025-145A-65 (1-618) x AAF73421 (1-1890)  
 QY 1 MetAlaLeuSerIleThrProLeuValSerArgSerCysLeuSer----- 16  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 60  
 QY 17 ---SerSerHisGluIleLysAlaLeuArgArgThrIleProThrLeuGlyIleCysArg 35  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 61  
 QY 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 121  
 QY 56 ---SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheille 74  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 181  
 QY 75 GlnSerLeuIleSerThrProTyrGlyValAlaProAspTyrArgGluArgAlaAspArgLeu 94  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 241  
 QY 95 IleGlyGluValLysAspIleMetPheAsnPhelLysSerLeuGluAspGlyGly----- 112  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 298  
 QY 113 -----AsnAspLeuGlnArgLeuLeuLeuValAspAspValGluArgLeuGly 129  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 354  
 QY 355 ACTCCCTCAATGATCTCATCAACGCCTTTTAAATGTCGATACGTTGAACGTTTAGGG 414  
 QY 130 IleAspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrp 149  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 149

Db 415 ATTGATAGACATTTCAAATAATGAGATAAAAGCAGCAGCTAGACTATGTTTACAGTTATTGG 474  
 QY 150 AsnGluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAla 169  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 475  
 QY 170 LeuGlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnVal 189  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 535  
 QY 190 PheLysAsp-----LysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGly 207  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 595  
 QY 208 GluIleArgGlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLys 227  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 652  
 QY 228 ValMetAspGluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIle 247  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 712  
 QY 248 ProAlaSerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTyrHisThr 267  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 772  
 QY 268 AsnLeuProArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsn 287  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 829  
 QY 288 LysAsnAla-----AlaGluLysLeuLeuGluAlaLysLeuGluPheAsnIle 304  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 889  
 QY 305 PheHisSerLeuGlnGluArgGluLeuLysHisValSerArgTyrTrpLysAspSerGly 324  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 949  
 QY 325 SerProGluMetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCys 344  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 1009  
 QY 345 IleAlaPheGluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeu 364  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 1069  
 QY 365 IleThrValLeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluPhe 384  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 1129  
 QY 385 ThrAlaThrIleLysArgTyrAspProSerAlaMetGluCysLeuProGluTyrMetLys 404  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 1189  
 QY 405 GlyValTyrMetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAla 424  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 1249  
 QY 425 GlnGlyArgAspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyr 444  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 1309  
 QY 445 MetGlnGluAlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGlu 464  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 1369  
 QY 465 AsnGlyLysValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAsp 484  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 1429  
 QY 485 IleProPheProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeu 504  
 Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 1489

QY 505 IleCysIleIleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArg 524  
 Db 1549 ATATGCTCCATCTTCGACTTAAATGACACTCGCTGCTACAGCGGACAGGCGCGT 1608  
 QY 525 GlyGluGluAlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGlu 544  
 Db 1609 GGAGAGAGAGCTTCGTATATCGTGTATATGAAGAGAACTCTGGATCAACAGAGNA 1668  
 QY 545 AspAlaLeuAsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGlu 564  
 Db 1669 GATGCTATCAATCAATCAACCTATGCTCAATACTAATCAAGAGAGTGAATGGGAG 1728  
 QY 565 LeuLeuLysProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSer 584  
 Db 1729 CTTCTCCGACAGGACGCGCCCTCATATGCTTGCAGAGAAACACGCTTTTGACATCCTC 1788  
 QY 585 ArgValTrpHisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThr 604  
 Db 1789 AAGAGTTCCTTCACGGCTACAAATACCGAGATGGTTCCAGCGTTGCCAACAGGAACC 1848  
 QY 605 LysSerLeuValMetArgThrValIleGluProValProLeu 618  
 Db 1849 AAGAATTGGGTGAGGAGACAGTCTCTGAGTCTGTGCTTTG 1890

RESULT 12

AAAF73413 standard; cDNA; 2429 BP.  
 AC AAF73413;  
 XX 30-APR-2001 (first entry)  
 DE Grand fir monoterpene synthase coding sequence fragment SEQ ID NO: 68.  
 KW Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
 KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
 KW terpinolene synthase; insect resistance; nutrition; ss.  
 OS Abies grandis.  
 XX WO200107565-A2.  
 XX 01-FEB-2001.  
 XX 24-JUL-2000; 2000WO-US020264.  
 XX 26-JUL-1999; 99US-00360545.  
 XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 XX Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
 XX WPI; 2001-182782/18.  
 XX P-PSDB; AAB69392.  
 XX New nucleic acid encoding monoterpene synthases, for increasing terpene  
 XX synthesis in plants, e.g. for increasing resistance to pests or for  
 XX treatment of cancer.  
 XX Claim 23; Page 156-159; 175pp; English.  
 XX The present invention provides the protein and coding sequences of  
 XX monoterpene synthases from the grand fir. These include (-)-camphene  
 XX synthase, (-)-beta-pinellene synthase, terpinolene synthase, (-)-  
 XX limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 XX and pinene synthase. The sequences can be used to produce transgenic  
 XX plants expressing high levels of the enzymes, resulting in levels which  
 XX are useful in protecting against and treating cancers, and to confer  
 XX insect resistance on plants  
 XX Sequence 2429 BP; 726 A; 484 C; 512 G; 707 T; 0 U; 0 Other;

Alignment Scores: 2,45e-205 Length: 2429  
 Pred. No.: 2084.50 Matches: 417  
 Score: 77.23% Conservative: 80  
 Best Local Similarity: 64.85% Mismatches: 115  
 Query Match: 64.12% Indels: 31  
 DB: 4 Gaps: 13  
 US-10-025-145a-65 (1-618) x AAF73413 (1-2429)  
 QY 1 MetAlaLeuLeuSerIleThrProLeu---ValSerArgSerCys----- 14  
 Db 35 ATGGCTCTCTCTCTATCTGATCTCTGAGGTTCCTCAAAATCTGGGGTGAATCGTTG 94  
 QY 15 LeuSerSerHisGluIleLysAlaLeu-----ArgArgThrIleProThrLeuGly 32  
 Db 95 ATCAGTTCAGCAATGTGCGAAGGCTCTCTGTATCTCTACAGCAGTCCCACTCTCAGA 154  
 QY 33 IleCysArgProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAla 52  
 Db 155 ATGCGTAGGCGACAGAAAGCTCTGTGTC-----ATCAACATGAAATTTGACCACTGTATCC 208  
 QY 53 SerThrAspSer-----ValGlnArgArgValGlyAsnTyrHisSerAsn 67  
 Db 209 CATCGTGATGATAATGCTGTGCTGCTACTGCAAGACGCATAGCCGATCATCATCCAAC 268  
 QY 68 LeuTrpAspAspPheIleGlnSerLeuIleSerThrProTyrGlyAlaProAspTyr 87  
 Db 269 CTGTGGAGAGATGATTTATCAATCATCTG---TCTCACCTTATGGGGATCTTCGTAC 325  
 QY 88 ArgGluArgAlaAspArgLeuIleGlyGluValLysAspIleMetPheAsnPhe----- 105  
 Db 326 AGTGAACGTGCTGTGACAGTGTGAGGAAGTAAAGAG---ATGTTCAATTCAATACCA 382  
 QY 106 -----LysSerLeuGluAspGlyGlyAsnLeuGluArgLeuLeuLeuVal 123  
 Db 383 AATAATAGAGAAATTTATTTGGTTCCCAAAATGATCTCTTACACGCTTTGGATGGTGGAT 442  
 QY 124 AspValGluArgGlyIleAspArgHisPheLysLysGluIleLysThrAlaLeuAsp 143  
 Db 443 AGCATTTGAACGTCGCGGATAGATAGACATTTCCAAAATGAGATAGAGTAGCCCTCGAT 502  
 QY 144 TyrValAsnSerTyrTrpAsnGluLys---GlyIleGlyCysGlyArgGluSerValVal 162  
 Db 503 TATGTTTACAGTTATTTGGAAGGAAAGAAAGCAATTTGGTGTGCGAGAGATTTACTTTT 562  
 QY 163 ThrAspLeuAsnSerThrAlaLeuGlyLeuArgThrLeuArgLeuHisGlyTyrThrVal 182  
 Db 563 CCTGATCTCACTCGACTCTCTGGCGCTTCGAACCTCTTCGACTGACGAGTACAATGTG 622  
 QY 183 SerSerAspValLeuAsnValPheLysAspLysAsnGlyGlnPheSerSerThrAlaAsn 202  
 Db 623 TCTTCAGATGTGCTGGAATCTTCAAGATCAAAAGGGGCATTTTCCTCCCTGCA--- 679  
 QY 203 IleGlnIleGluGlyGluIle---ArgGlyValLeuAsnLeuPheArgAlaSerLeuVal 221  
 Db 680 ATCTTAACCGGACAGACAGATCACTAGAGTGTCTTAAATTTATATATCGGCTTCCCTGGTC 739  
 QY 222 AlaPheProGlyGluLysValMetAspGluAlaGluThrPheSerThrLysTyrLeuArg 241  
 Db 740 GCCTTTCCGGGGAGAAAGTTATGGAAGAGGCTGAAATCTTCTCGGCATCTTATTTGAAA 799  
 QY 242 GluAlaLeuGlnLysIleProAlaSerSerIleLeuSerLeuGluIleArgAspValLeu 261  
 Db 800 GAAGTCTTACAAAGATTCAGTCTCCAGT---TTTTACGAGAGATAGATAACGTTTGT 856  
 QY 262 GluTyrGlyTrpHisThrAsnLeuProArgLeuGluAlaArgAsnTyrMetAspValPhe 281  
 Db 857 GAATATGTTGGCACACAAATTTGCCAAGATTTGGAAGCAAGAAATTTATATCGACGCTTAC 916  
 QY 282 GlyGlnHisThrLysAsnLysAsn-----AlaAlaGluLysLeuLeu 295  
 Db 917 GGCAGGACAGCTATGAAAGTTCAACAGAGATGCCATATGTGNATACGAGAGACTTTTA 976

```

QY 296 GluLeuAlaLysLeuGluPheAsnIlePheHisSerLeuGlnGluArgGluLeuLysHis 315
Db 977 AAACCTGCAAAATGGAGTTTAATATCTTCACTCTTTGCACAGAAAGAGTTGCAATAT 1036
QY 316 ValSerArgTTPTrpLysAspSerGlySerProGluMetThrPheCysArgHisArgHis 335
Db 1037 ATCTCTAGATGGTGAAGATTCGTGTCACTCATCTGACCTTTTACTCGACACCGTCCAC 1096
QY 336 ValGluTyrTyrAlaLeuAlaSerCysIleAlaPheGluProGlnHisSerGlyPheArg 355
Db 1097 GTGGAATACTACAAATGGCATCTTGCATTTCTATGGAGCGGAACACTCCGCTTTCAGA 1156
QY 356 LeuGlyPheThrLysMetSerHisLeullethrValLeuAspAspMetTyrAspValPhe 375
Db 1157 TTGGGGTTTGTCAAAACATGTCATCTCTAACAGTTCTGGATGATATATGACACTTTT 1216
QY 376 GlyThrValAspGluLeuGluPheThrAlaThrIleLysArgTTPAspProSerAla 395
Db 1217 GGAACACTGGAGCACTCCACTTTTACGACTGCTTTAAGAGATGGATTTGTCAGAG 1276
QY 396 MetGluCysLeuProGluTyrMetLysGlyValTyrMetMetValTyrHisThrValAsn 415
Db 1277 ACAAGTGTCTTCCAGATATATGAAGCAGTGTACATGGACTTGTATCAATGTCTTAAT 1336
QY 416 GluMetAlaArgValAlaGluLysAlaGlnGlyArgAspThrLeuAsnTyrAlaArgGln 435
Db 1337 GAATGGCGCAAGAGCTGGAAGACTCAAGCGAGAGATACGCTCAACTATATTCGCAAT 1396
QY 436 AlaTrpGluAlaCysPheAspSerTyrMetGlnGluAlaLysTTPAlaThrGlyTyr 455
Db 1397 GCTTATGAGTCTCATTTTGTATTCGTTTATGCACGACAAATGGAUCTCAAGTGGTAT 1456
QY 456 LeuProThrPheGluGluTyrLeuGluAsnGlyLysValSerSerAlaHisArgProCys 475
Db 1457 CTCCCAAGTGTGAGGAGTCTTGAAGAATGGAAAGTTAGTTCGGTTCCTCGCACGCC 1516
QY 476 AlaLeuGlnProIleLeuThrLeuAspIleProPheProAspHisIleLeuLysGluVal 495
Db 1517 ACTTTACACCACTACTCACCCTGGATGTACCATCTTCCATAATTACATCTCAAGAAAT 1576
QY 496 AspPheProSerLysLeuAsnAspIleCysIleLeuArgLeuArgGlyAspThr 515
Db 1577 GATATCACTAGTTCATGACTTGGCTTCGCTCCCTCCTTCGGTACGTACGTGGTACAG 1636
QY 516 ArgCysTyrLysAlaAspArgAlaArgGlyGluGluAlaSerSerIleSerCysTyrMet 535
Db 1637 CGCTGCTACAAAGCGGATAGGCGTCTGGAGAAGAAGCTTCAGCTATATCGTTATATG 1696
QY 536 LysAspAsnProGlyLeuThrGluAspAlaLeuAsnHisIleAsnPheMetIleArg 555
Db 1697 AAAGACCATCTCGATCAACAGAGGAAGATCTCTCAATCATATCAACGTCATGATCAGT 1756
QY 556 AspAlaIleArgGluLeuAsnTrpGluLeuLysProAspAsnSerValProIleThr 575
Db 1757 GATCCAAATCAGAAATTAATTTGGAGCTTCTCAGACCAGATAGCAAAAGTCCCATCTCT 1816
QY 576 SerLysLysHisAlaPheAspIleSerArgValTrpHisHisGlyTyrArgTyrArgAsp 595
Db 1817 TCCAAAGAAACATGCTTTTGACATCACCAGAGCTTTCCATCACCTCTCAAGTACCGAGAT 1876
QY 596 GlyTyrSerPheAlaAsnValGluThrLysSerLeuValMetArgThrValIleGluPro 615
Db 1877 GGTTACATGTTGCGAGTAGTGAACAAAGAAATTTGGTGATGAAGAAACAGTTCCTTGAACCT 1936
QY 616 ValProLeu 618
Db 1937 GTGGCATTG 1945

```

RESULT 13

AAAX08645

ID AAAX08645 standard; cDNA; 2089 BP.

XX

```

AC AAAX08645;
XX 27-SEP-1999 (first entry)
XX Limonene synthase gene.
DE Myrcene synthase; limonene synthase; pinene synthase; flavour; aroma;
KW defense; plant seed; oil; meal; ss.
XX Abies grandis.
XX Key Location/Qualifiers
XX CDS 73..1986
XX /*tag= a
XX /product= "Limonene synthase"
PN WO9902030-A1.
XX 21-JAN-1999.
XX 10-JUL-1998; 98WO-US014528.
XX 11-JUL-1997; 97US-0052249P.
XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX Bohlmann J, Steele CL, Croteau RB;
XX WPI; 1999-120396/10.
XX P-PSDB; AAW85702.
XX New isolated gymnosperm monoterpene synthase DNA - obtained from Grand
XX fir (Abies grandis), used to provide plants with modified production of
XX monoterpenes, e.g. myrcene, limonene or pinene.
XX Claim 11; Page 79-82; 121pp; English.
XX Nucleotide sequences encoding myrcene synthase, limonene synthase and
XX pinene synthase from Grand fir may be incorporated into any organism
XX (e.g. intact plant, animal, microbe), or derived cell culture that
XX produces geranyl diphosphate for the production of the aforementioned
XX enzymes or their products. The sequences when expressed in transfected
XX cells may also be used for the production or modification of flavour and
XX aroma properties, improvement of defense capability, and the alteration
XX of other ecological interactions mediated by myrcene, limonene, pinene,
XX or their derivatives. In particular they can be used for the production
XX of plant seeds for the extraction of oil or meal

```

SQ Sequence 2089 BP; 538 A; 415 C; 450 G; 586 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	6,08e-200	Length:	2089
Score:	2031.50	Matches:	409
Percent Similarity:	76.67%	Conservative:	84
Best Local Similarity:	63.61%	Mismatches:	119
Query Match:	62.49%	Indels:	31
DB:	2	Gaps:	14

US-10-025-145A-65 (1-618) x AAAX08645 (1-2089)

```

QY 1 MetAlaLeuLeuSerIleThrProLeu---ValSerArgSerCys----- 14
Db 73 ATGGCTCTCCTTTCTATCGTATCTTTTCAGAGTTCCCAAAATCCTCGCGGCTGAAATCGTTG 132
QY 15 LeuSerSerSerHisGluIleLysAlaLeu-----ArgArgThrIleProThrLeuGly 32
Db 133 ATCAGTTCACGCAATGTGCAGAGGCTCTCTGTATCTCTACAGCAGTCCCAACACTCAGA 192
QY 33 IleCysArgProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAla 52
Db 193 ATGCGTAGGCGACAGAAAGCTCTGGTC-----ATCACATGAAATGACCCTGTATCC 246
QY 53 SerThrAspSer-----ValGlnArgArgValGlyAsnTyrHisSerAsn 67

```



PR 22-APR-1999; 99US-0130628P.  
 PR 23-AUG-1999; 99US-0150262P.  
 PA (KENT ) UNIV KENTUCKY RES DEPT.  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX Chappell J, Manna KR, Noel JP, Starks CM;  
 XX WPI; 2000-292839/25.  
 XX P-PSDB; AAY90859.  
 XX Novel terpene synthase enzymes, useful for producing terpene  
 PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known  
 PT enzymes by specific amino acid alterations.  
 XX  
 PS Disclosure; Page 445-448; 450pp; English.  
 XX  
 CC The present invention describes an isolated terpene synthase (I)  
 CC comprising a region with at least 20% identity to region 265-535 of a 548  
 CC amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha  
 CC -carbon atoms (alphaC) that have interatomic distances, between each  
 CC other, within tabulated ranges, have a centre point (within a sphere of  
 CC radius 2.3 Angstrom) within tabulated ranges, and have an ordered  
 CC arrangement of R groups (defining aa side chains), excluding specific  
 CC tabulated arrangements (tables given in the specification). (I), and  
 CC related enzymes, are used to produce a wide range of terpenoids (e.g.  
 CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,  
 CC flavours, pheromones, defensive agents, pigments, antitumour agents,  
 CC components of signal transduction pathways, precursors of steroid  
 CC hormones and bile acids, as photoreceptors and as co-factor side  
 CC chains.  
 CC Some synthases with little or no catalytic activity (and nucleic acids  
 CC encoding them) are used as controls in the analysis of products formed by  
 CC enzymatic synthesis; as nutrient supplements; for affinity purification  
 CC of isoprenoids; or to develop immunological reagents or nucleic acids for  
 CC monitoring expression of terpene synthase or inheritance of the gene in  
 CC plant breeding programs. The new synthases may produce novel terpene  
 CC products. AAY90831 to AAY90838 and AAY90831 to AAY90859 represent  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 2089 BP; 638 A; 415 C; 450 G; 586 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 5,08e-200 Length: 2089  
 Score: 2031.50 Matches: 409  
 Percent Similarity: 76.67% Conservative: 84  
 Best Local Similarity: 63.61% Mismatches: 119  
 Query Match: 62.49% Indels: 31  
 DB: 3 Gaps: 14  
 US-10-025-145A-65 (1-618) x AAY90838 (1-2089)  
 Qy 1 MetAlaLeuSerIleThrProLeu---ValSerArgSerCys----- 14  
 Db 73 ATGGCTCTCCCTTCTATCGTATCTTTGCGAGGTCCCAAAATCTCGGCGCTGAAATCGTTG 132  
 Qy 15 LeuSerSerHisGluIleGlyAlaLeu-----ArgArgThrIleProThrLeuGly 32  
 Db 133 ATCAGTTCCAGCAATGTGCAAGAGGCTCTCTGTATCTCTACAGCAGTCCCAACACTCAGA 192  
 Qy 33 IleCysArgProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAla 52  
 Db 193 ATGGTAGGCAACAGAAAGCTCTGGTC-----ATCAACATGAATTTGACCACCTGTATCC 246  
 Qy 53 SerThrAspSer-----ValGlnArgArgValGlyAsnTyrHisSerAsn 67  
 Db 247 CATCGTGATGATAATGGTGGTGTACTGCAAGAGCGCATACCGCATCATCCCCAAC 306  
 Qy 68 LeuTrpAspAspPheIleGlnSerLeuIleSerThrProTyrGlyAlaProAspTyr 87  
 Db 307 CTGTGGGAAGATGATTTTCATACATCATTTG---TCCTCACCTTATGGGGGATCTTCGTAC 363  
 Qy 88 ArgGluArgAlaAspArgLeuIleGlyGluValLysAspIleMetPheAsnPhe----- 105

Db 364 AGTGAACGTGCTGACACAGTCGTGTGAGGAAGATAAAGAG---ATGTTCAATTCAATACCA 420  
 Qy 106 -----LysSerLeuGluAspGlyGlyAsnAspLeuLeuGlnArgLeuLeuValAsp 123  
 Db 421 AATAATAGAGAATTATTTGGTTCCCAAAATGATCTCTTACACGCTTGGATGGTGGAT 480  
 Qy 124 AspValGluArgLeuGlyIleAspArgHisPheLysLysGluIleLysThrAlaLeuAsp 143  
 Db 481 AGCATGAACGCTCTGGGATAGATAGATTCATTTCCAAAATGAGATAAGAGTAGACCCCTCAT 540  
 Qy 144 TyrValAsnSerTyrTrpAsnGluLys---GlyIleGlyCysGlyArgGluSerValVal 162  
 Db 541 TATGTTTACAGTTATTTGGAAGAAAGAGGAGCATTTGGGTGTGGCAGAGATTTCTACTTTT 600  
 Qy 163 ThrAspLeuAsnSerThrAlaLeuGlyLeuArgThrLeuArgLeuHisGlyTyrThrVal 182  
 Db 601 CCTGATCTCAACTCGACTGCTTGGCGCTTGGAACTCTTCGACTGCAGGATACATATGTG 660  
 Qy 183 SerSerAspValLeuAsnValPheLysAspLysAsnGlyGlnPheSerSerThrAlaAsn 202  
 Db 661 TCTTCAGATGTCTGGAATACCTTCAAAAGATGAAAAGGGGCAATTTTGCCTGCCCTGCA--- 717  
 Qy 203 IleGlnIleGluGlyGluLeu---ArgGlyValLeuAsnLeuPheArgAlaSerLeuVal 221  
 Db 718 ATCTTAACCGAGGACAGATCCTAGAAAGTGTCTTAAATTTATATCGGGCTTCCCTGGTC 777  
 Qy 222 AlaPheProGlyGluLysValMetAspGluAlaGluThrPheSerThrLysTyrLeuArg 241  
 Db 778 GCCTTTCCCGGGGAGAAAGTTATGGAAGAGGCTGAAATCTTCTCGCATCTTATTGAA 837  
 Qy 242 GluAlaLeuGlnLysIleProAlaSerSerIleLeuSerLeuGluIleArgAspValLeu 261  
 Db 838 AAGCTCTTACAAAAGATTCGCGTCTCCAAT---CTTTCAGGAGAGATAGATAATATGTTT 894  
 Qy 262 GluTyrGlyTyrHisThrAsnLeuProArgLeuGluAlaArgAsnTyrMetAspValPhe 281  
 Db 895 GAATATGTTGGCACACGAATTTGCCGAGATTGGAAGCAAGAAATTTATCAGGCTCTAC 954  
 Qy 282 -----GlyGlnHisThrLysAsnLysAla-----AlaGluLysLeuLeu 295  
 Db 955 GAGCAGAGCGGCTATGAAAGCTTAAACGAGATGCCATATATGAACATCAAGAGCTTTA 1014  
 Qy 296 GluLeuAlaLysLeuGluPheAsnIlePheHisSerLeuGlnGluArgGluLeuLysHis 315  
 Db 1015 CAACTTGCAAAATTGGAGTTTCATATCTTTCACCTTTTGCACCTAAGAGAGATTACATCT 1074  
 Qy 316 ValSerArgTrpTrpLysAspSerGlySerProGluMetThrPheCysArgHisArgHis 335  
 Db 1075 ATCTCCAGATGGTGAAGAAATCAGGTTCTCACTGACTTTTACACGGCATCGTCAC 1134  
 Qy 336 ValGluTyrTyrAlaLeuAlaSerCysIleAlaPheGluProGlnHisSerGlyPheArg 355  
 Db 1135 GTGGAATACTACACTATGGCATCTTCATTTCTATGTTGCCAAACATTCAGCTTCAGA 1194  
 Qy 356 LeuGlyPheThrLysMetSerHisLeuIleThrValLeuAspAspMetTyrAspValPhe 375  
 Db 1195 ATGGAGTTTGTCAAAGTGTGTCTTCTTAACAGTTCTCGATGATATATATGACACTTTT 1254  
 Qy 376 GlyThrValAspGluLeuGluPheThrAlaThrIleLysArgTrpAspProSerAla 395  
 Db 1255 GGAACATGAAACGAACTTCCAACTTTTACGGATGCAATTAAGAGATGGGATTTGTCAACG 1314  
 Qy 396 MetGluCysLeuProGluTyrMetLysGlyValTyrMetMetValTyrHisThrValAsn 415  
 Db 1315 ACAAGTGGCTTCCAGAAATATGAAAGAGTGTACATGGACTGTATCAATGATTAAT 1374  
 Qy 416 GluMetAlaArgValAlaGluLysAlaGlnGlyArgAspThrLeuAsnTyrAlaArgGln 435  
 Db 1375 GAAATGGTGAAGAGGCTGAGAGACTCAAGCGCGAGATATGCTCAATATATCAAAAT 1434  
 Qy 436 AlaTrpGluAlaCysPheAspSerTyrMetGlnGluAlaLysTrpIleAlaThrGlyTyr 455  
 Db 1435 GCITGGGAAGCCCTATTGTATACCTTTATGCAAGCAAGTGGATCTCCAGCAGTTAT 1494

QY 456 LeuProThrPheGluThrLeuGluAsnGlyLysValSerSerAlaHisArgProCys 475  
 Db 1495 CCCCACCGTTGAGGAGTACITGAAGATGCAAAAGTAGTTCTGGTCTCCATAGCC 1554  
 QY 476 AlaLeuGlnProIleLeuThrLeuAspIleProPheProAspHisIleLeuLysGluVal 495  
 Db 1555 ACATTACACCCATTCTCACTTGGATGATACCACTTCTCTGATTACATCTGCAAGAAAT 1614  
 QY 496 AspPheProSerLysLeuAsnAspLeuIleCysIleIleLeuArgLeuArgGlyAspThr 515  
 Db 1615 GATTATCCATCCAGATTCAATGATTAGCTTCGTCCATCTCTGCTACCTAGAGTGACAG 1674  
 QY 516 ArgCysTyrLysAlaAspArgAlaArgGlyGluGluAlaSerSerIleSerCysTyrMet 535  
 Db 1675 CGCTGTCTACAGCGGATAGGCGCGGTGGAGAAAGCTTCAGCTATATCTGTTTAATG 1734  
 QY 536 LysAspAsnProGlyLeuThrGluGluAspAlaLeuAsnHisIleAsnPheMetIleArg 555  
 Db 1735 AAGACCATCTGGATCAATAGAGGAAGATGCTCTCAATCATATCAACGCCATGATCAGT 1794  
 QY 556 AspAlaIleArgGluLeuAsnTrpGluLeuLeuLysProAspAsnSerValProIleThr 575  
 Db 1795 GATGCAATCAGAAATTAATTTGGAGCTTCTCAGACCGGATAGCAAAAGTCCCATCTCT 1854  
 QY 576 SerLysLysHisAlaPheAspIleSerArgValTrpHisGlyTyrArgTyrArgAsp 595  
 Db 1855 TCCAGAAACATGCTTTTGACATCACCAGAGCTTCCATCATGCTTACAAATATCAGAT 1914  
 QY 596 GlyTyrSerPheAlaAsnValGluThrLysSerLeuValMetArgThrValIleGluPro 615  
 Db 1915 GGTTACACTGTTTCCACACGAAACAAAGAAATTTGGTGTGATGAACCGTTCTTGAACCT 1974  
 QY 616 ValProLeu 618  
 Db 1975 CTGCTTTG 1983

RESULT 15

AAAF73373  
 ID AAF73373 standard; cDNA; 2089 BP.

AC AAF73373;

XX 30-APR-2001 (first entry)

DE Grand fir (-)-limonene synthase coding sequence SEQ ID NO: 5.

XX Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
 KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
 KW terpinolene synthase; insect resistance; nutrition; ss.

XX Abies grandis.

XX WO200107565-A2.

XX 01-FEB-2001.

XX 24-JUL-2000; 2000WO-US020264.

XX 26-JUL-1999; 99US-00360545.

XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX Steele CL, Bohlmann J, Croteau RB, Phillips MA;

XX WPI; 2001-182782/18.

XX P-PSDB; AAB69372.

XX New nucleic acid encoding monoterpene synthases, for increasing terpene  
 PT synthesis in plants, e.g. for increasing resistance to pests or for  
 PT treatment of cancer.

XX Claim 33; Page 112-115; 175pp; English.

XX The present invention provides the protein and coding sequences of  
 CC monoterpene synthases from the grand fir. These include (-)-camphene  
 CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 CC limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 CC and pinene synthase. The sequences can be used to produce transgenic  
 CC plants expressing high levels of the enzymes, resulting in levels which  
 CC are useful in protecting against and treating cancers, and to confer  
 CC insect resistance on plants

XX SQ Sequence 2089 BP; 638 A; 415 C; 450 G; 586 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6,08e-200 Length: 2089  
 Score: 2031.50 Matches: 409  
 Percent Similarity: 76.67% Conservative: 84  
 Best Local Similarity: 63.61% Mismatches: 119  
 Query Match: 62.49% Indels: 31  
 DB: 4 Gaps: 14

US-10-025-145A-65 (1-618) x AAF73373 (1-2089)

QY 1 MetAlaLeuSerIleThrProLeu---ValSerArgSerCys----- 14  
 Db 73 ATGGCTCTCTTCTATCTGATCTTTGCGAGTTCCCAATCTCTGGGGCTGAAATCGTTG 132  
 QY 15 LeuSerSerSerHisGluIleLysAlaLeu-----ArgArgThrIleProThrLeuGly 32  
 Db 133 ATCAGTTCAGCAATGTGCGAAGGCTCTCTGATCTCTACAGCAGTCCCAACACCTCAGA 192  
 QY 33 IleCysArgProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAla 52  
 Db 193 ATGCGTAGCGCAGAGAAGCTCTGGTC-----ATCAACATGAATGACCATGATCC 246  
 QY 53 SerThrAspSer-----ValGlnArgValGlyAsnTyrHisSerAsn 67  
 Db 247 CATCGTGATGATAAATGGTGGTGTGCTGCTGCTGCAAGACGATATCCGATATCCCAAC 306  
 QY 68 LeuTrpAspAspPheIleGlnSerLeuIleSerThrProTyrGlyAlaProAspTyr 87  
 Db 307 CTGTGGGAAGATGATTTCATACATCATTTG---TCCTCACCTTATGGGGATCTTCGTAC 363  
 QY 88 ArgGluArgAlaAspArgLeuIleGlyGluValLysAspIleMetPheAsnPhe----- 105  
 Db 364 AGTGAACGTGCTGAGACAGTCTGTTGAGGAAGTAAAGAG---ATGTTCAATTCAATACCA 420  
 QY 106 -----LysSerLeuGluAspGlyGlyAsnAspLeuLeuGlnArgLeuLeuValAsp 123  
 Db 421 AATAATAGAGAATTAATTTGGTTCCCAAAATGATCTCTTACAGCCCTTTGGATGGTGGAT 480  
 QY 124 AspValGluArgLeuGlyIleAspArgHisPheLysLysLysLysLysThrAlaLeuAsp 143  
 Db 481 AGCATTTGAACGCTCTGGGGATAGATAGATCATTTCCAAATGAGATAGATAGAGCTTCGAT 540  
 QY 144 TyrValAsnSerTyrTrpAsnGluLys---GlyIleGlyCysGlyArgGluSerValVal 162  
 Db 541 TATGTTTACAGTTATTGGAAGGAAAGGAAGGAGGATTTGGGTGTGGCAGAGATTTCTACTTT 600  
 QY 163 ThrAspLeuAsnSerThrAlaLeuGlyLeuArgThrLeuArgLeuHisGlyTyrThrVal 182  
 Db 601 CTGATCTCAACTCGACTGCTTGGCGCTTCGAACTTCGAACTTCGACGCGATACAATGTG 660  
 QY 183 SerSerAspValLeuAsnValPheLysAspLysAsnGlyGlnPheSerSerThrAlaAsn 202  
 Db 661 TCTTCAGATGTCTGGAATCTTCAAGATGAAAGGGGCAATTTTGGCTGCCCTGCA--- 717  
 QY 203 IleGlnIleGluGlyGluLe---ArgGlyValLeuAsnLeuPheArgAlaSerLeuVal 221  
 Db 718 ATCCTAACCGAGGAGCAGATCACTAGAGTGTCTTAAATTTATATCGGGCTTCCCTGGTC 777  
 QY 222 AlaPheProGlyGlyLysValMetAspGluAlaGluThrPheSerThrLysTyrLeuArg 241  
 Db 778 GCCTTTCCCGGGGAGAAAGTTATGGAAGAGGCTGAAATCTTCTCGGCACTTATTATGAAA 837



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 8, 2004, 15:05:55 ; Search time 4324 Seconds  
(without alignments)  
4267.999 Million cell updates/sec

Title: US-10-025-145A-65  
Perfect score: 3251  
Sequence: 1 MALLSITPLVRSCLSSSHE.....FANVETKSLVMTVIEPVPL 618

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US10025145/runat\_07072004\_160150\_27949/app\_query.fasta\_1.775  
-DB=EST -QFMT=Fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10025145.cgn2\_1\_3437@runat\_07072004\_160150\_27949 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1066	32.8	745	14	CF477103	CF477103 RTW3_5 A
2	907	27.9	700	14	CF474786	CF474786 RTW2_7 B
3	888	27.3	696	14	CF401916	CF401916 RTW1_15
4	872	26.8	669	14	CF479802	CF479802 RTW3_12
5	864	26.6	697	13	BQ196773	BQ196773 NXLV105_B
6	853	26.2	599	13	BX680641	BX680641 RTW3_12
7	802	24.7	599	12	BG526917	BG526917 NXPV 057
8	795	24.5	804	14	CF666338	CF666338 RTCNT1_22
9	786	24.2	616	14	CF663845	CF663845 RTCNT1_5
10	764	23.5	637	14	CF476978	CF476978 RTW3_5_A
11	733	22.5	481	9	AL750951	AL750951 AL750951
12	723	22.2	740	14	CF477562	CF477562 RTW3_8 G
13	717	22.1	516	13	BQ698077	BQ698077 NXPV 064
14	713	21.9	517	13	BX677624	BX677624 RTW3_12
15	698	21.5	539	10	AW287756	AW287756 EST0004 S
16	681	20.9	569	10	AW287754	AW287754 EST0002 S
17	680	20.9	557	14	CF666483	CF666483 RTCNT1_23
18	658	20.2	651	14	CF479884	CF479884 RTW3_12
19	657.5	20.2	534	10	AW287755	AW287755 EST0003 S
20	656	20.2	682	14	CF663768	CF663768 RTCNT1_5
21	607	18.7	430	13	BQ702557	BQ702557 NXSI 129
22	587.5	18.1	591	14	CA305371	CA305371 hasp004xxk
23	579.5	17.8	2029	11	AY107369	AY107369 Zea mays
24	559	17.2	618	9	AL750955	AL750955 AL750955
25	550	16.9	541	14	CF672979	CF672979 RTCNT1_75
26	526.5	16.2	725	14	CF397293	CF397293 RTDS3_2 A
27	525	16.1	542	10	BG039521	BG039521 NXSI 099
28	517.5	15.9	1981	11	AY105371	AY105371 Zea mays
29	515.5	15.9	725	14	CF470471	CF470471 RTD31_17
30	507.5	15.6	629	9	AU299287	AU299287 AU299287
31	496.5	15.3	843	14	CF208265	CF208265 CAB20003
32	489	15.0	574	14	CF474640	CF474640 RTW2_7 B
33	489	15.0	733	14	CF666270	CF666270 RTCNT1_22
34	488	15.0	2028	11	AY110842	AY110842 Zea mays
35	481	14.8	730	13	BX682869	BX682869 BX682869
36	477.5	14.7	896	29	CG847800	CG847800 ZMMBB031
37	476	14.6	634	14	CF397946	CF397946 RTDS3_23
38	475	14.6	597	14	CF666416	CF666416 RTCNT1_23
39	465.5	14.3	585	12	BG318038	BG318038 NXPV 008
40	453	13.9	574	9	AW065088	AW065088 ST39E0A P
41	446.5	13.7	766	13	BUR14724	BUR14724 N032C12 P
42	439.5	13.5	552	13	BQ696036	BQ696036 NXPV 036
43	429.5	13.2	530	9	AU298829	AU298829 AU298829
44	427	13.1	760	12	BG351484	BG351484 111F08 Ma
45	418.5	12.9	824	14	CD852160	CD852160 DH0ALL25Z

# ALIGNMENTS

RESULT 1  
CF477103  
LOCUS  
DEFINITION RTW3\_5\_A06.g1\_A022 Well-watered loblolly pine roots W33 Pinus  
ACCESSION CF477103  
VERSION CF477103  
KEYWORDS EST.  
ORGANISM Pinus taeda (loblolly pine)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
REFERENCE 1 (bases 1 to 745)



(Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 9,79e-94 Length: 700  
 Score: 907.00 Matches: 160  
 Percent Similarity: 84.98% Conservative: 38  
 Best Local Similarity: 68.67% Mismatches: 35  
 Query Match: 27.90% Indels: 0  
 DB: 14 Gaps: 0

US-10-025-145A-65 (1-618) x CF474786 (1-700)

QY 334 ArgHisValGluTyrValAlaLeuAlaSerCysIleAlaPheGluProGlnHisSerGly 353  
 Db 2 CGTCAGGTGGATACATACATTTAGCAGCTTGCAATGGATGCAATGATCTTAAACATTTCTCGG 61  
 QY 354 PheArgLeuGlyPheThrLysMetSerHisLeuIleThrValLeuAspMetTyrAsp 373  
 Db 62 TTTCGACTAGGATTTGGTAAATAAGTCATATGATCAGATTCTCGACGATATCTACGAC 121  
 QY 374 ValPheGlyThrValAspGluLeuGluPheThrAlaThrIleLysArgTirAspPro 393  
 Db 122 ACCITCGGAACATGAGGAGCTCGAATCTTTAACCGCAGCGGTTTAAAGAGATGGGATCGG 181  
 QY 394 SerAlaMetGluCysLeuProGluTyrMetLysGlyValTyrMetMetValTyrHisThr 413  
 Db 182 TCTTCGATAGAGTGTCTTCCAGATTATATGAAGAGGTGTATATGCGGTTTACGACAC 241  
 QY 414 ValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArgAspThrLeuAsnTyrAla 433  
 Db 242 ATCAACGAAATGGCAGCAGCGCGAGAAATTCGAAGCTGGGATACAGTCAGTCATGCT 301  
 QY 434 ArgGlnAlaTrrGluAlaCysPheAspSerTyrMetGlnGluAlaLysTrrIleAlaThr 453  
 Db 302 CGAAATCTTGGGAGGCTTTTATTTGGTCTTATATACAAAGAGCCAACTGGATTTCCAGT 361  
 QY 454 GlyTyrLeuProThrPheGluTyrLeuGluAsnGlyLysValSerSerAlaHisArg 473  
 Db 362 GGTATCTTCCACGTTTCGACGAGTACCTCGAGATGGAGAGTTCAGCTCGGCTCTCGC 421  
 QY 474 ProCysAlaLeuGlnProIleLeuThrLeuAspIleProPheProAspHisIleLeuLys 493  
 Db 422 ATAACCCAGCTCGAACCCATCTGACTTGGGTTTCTCTTCGCGCTCGAATCTTCGAC 481  
 QY 494 GluValAspPheProSerLysLeuAsnAspLeuIleCysIleIleLeuArgLeuArgGly 513  
 Db 482 GAATTCGACTTTCATCGAATTCATGATTGTATGTGTCATCTTCGACTGAAAGGT 541  
 QY 514 AspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGluAlaSerSerIleSerCys 533  
 Db 542 GACACTCAATGCTACAGGCTGACAGGGCGGTGGAGAGAGCTTTCGGCGGTACGTGT 601  
 QY 534 TyrMetLysAspAsnProGlyLeuThrGluAspAlaLeuAsnHisIleAsnPheMet 553  
 Db 602 TATATGAAGACCATCTCGAATAACAGAGAGATGCTGTCAATCAATCAATGCTATG 661  
 QY 554 IleArgAspAlaIleArgGluLeuAsnTrpGluLeuLeu 566  
 Db 662 GTCGATAACTTAACCAAGGAAGTGAATGGAGGTACTT 700

## RESULT 3

CF401916

LOCUS

DEFINITION

RTW1\_15\_B05\_g1\_A015 Well-watered loblolly pine roots WW1 Pinus

taeda\_cDNA clone RTW1\_15\_B05\_A015 5', mRNA sequence.

CF401916

CF401916.1 GI:34360333

KEYWORDS

SOURCE  
ORGANISM

Pinus taeda (loblolly pine)

Pinus taeda

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

1 (bases 1 to 696)

Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J., Gebremedhin, M., Dervin, C., Martin, T., White, T., Davis, J., and Neale, D.

An EST database from well-watered loblolly pine (Pinus taeda) roots

Unpublished (2003)

Contact: Cordonnier-Pratt, M.M.

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.

Seq primer: JENREV (CAGGAAACAGCTATGACC).

Location/Qualifiers

1..696

/organism="Pinus taeda"

/mol\_type="mRNA"

/strain="CCONES"

/db\_xref="taxon:3352"

/clones="RTW1\_15\_B05\_A015"

/lab\_host="DH10B-T1 Phage-resistant E. coli"

/clone\_lib="Well-watered loblolly pine roots WW1"

/note="Vector: pSL180; Site1: EcoRI; Site2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

FEATURES

source

## ORIGIN

Alignment Scores:

Pred. No.: 1.53e-91 Length: 696  
 Score: 888.00 Matches: 156  
 Percent Similarity: 84.48% Conservative: 40  
 Best Local Similarity: 67.24% Mismatches: 36  
 Query Match: 27.31% Indels: 0  
 DB: 14 Gaps: 0

US-10-025-145A-65 (1-618) x CF401916 (1-696)

QY 344 CysIleAlaPheGluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHis 363  
 Db 1 TGCATTGCAATGATCTCTAAACATTTTTCGCTTCGACTAGGATTTGCTAAATAAGTCAT 60  
 QY 364 LeuIleThrValLeuAspMetTyrAspValPheGlyThrValAspGluLeuGluLeu 383  
 Db 61 ATGATCAGGATTTTCGACGATATCTACGACCTTCGGAACAATGGAGGAGCTCGAATC 120  
 QY 384 PheThrAlaThrIleLysArgTrrAspProSerAlaMetGluCysLeuProGluTyrMet 403  
 Db 121 TTAACCGCAGCGTTTAAAGATGGATCGCTCTTCGATAGAGTGTCTTCAGATTATATG 180  
 QY 404 LysGlyValTyrMetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLys 423  
 Db 181 AAAGGAGTGTACATGCGGTTTACGACACATCAACGAAATGGCAGAGGCGCAGAAA 240  
 QY 424 AlaGlnGlyArgAspThrLeuAsnTyrAlaArgGlnAlaTrrGluAlaCysPheAspSer 443  
 Db 241 ATTCAAGGCTGGGATACAGTCAGCTATGCTCGAAATCTTTGGAGGCTTTTATTGGTCT 300



```

ACCESSION BQ196773
VERSION BQ196773.1 GI:20379276
SOURCE EST.
KEYWORDS Pinus taeda (loblolly pine)
ORGANISM Pinus taeda
REFERENCE Molecular Basis of Wood Formation in the Pine Megagenome
AUTHORS Sederoff, R.
TITLE Unpublished (2000)
JOURNAL Forest Biotechnology
COMMENT Contact: Sederoff, Ron
North Carolina State University
840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,
NC 27695, USA
Tel: 919 515 7800
Fax: 919 515 7801
Email: ron.sederoff@ncsu.edu, jerri.johnson@ncsu.edu
Please see http://web.ahc.umn.edu/biodata/nsfpine/ for further
information.
Seq primer: T3.
FEATURES             Location/Qualifiers
     source
       1..697
         /organism="Pinus taeda"
         /mol_type="mRNA"
         /strains="Coastal plain loblolly pine from North Carolina"
         /db_xref="taxon:3352"
         /clone="NXLIV105 B02"
         /tissue_type="primary xylem"
         /dev_stage="late wood"
         /lab_host="XLI-Blue"
         /clone_lib="NXLI-Blue"
         /note="Vector: pTriplex; Site 1: EcoRI; The library is
from late (summer-August) wood, taken from below the crown
of a 20 year old tree. The harvested xylem tissue was on
the the cusp between transitional and mature wood. NOTE:
The sequences contain a 'cDNA adapter' between the EcoRI
site and the start of the EST. The adapter sequence is
'AATTCGCCATTATGGCC'."
ORIGIN
Alignment Scores:
Pred. No.:          9,218-89          Length:          697
Score:             864.00             Matches:         162
Percent Similarity: 85.45%             Conservative:    20
Best Local Similarity: 76.06%           Mismatches:      31
Query Match:       26.58%             Indels:          0
DB:                13                 Gaps:            0

US-10-025-145A-65 (1-618) x BQ196773 (1-697)
Qy 406 ValTyrMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGln 425
Db 23 GTTTACATGATAGTTTACACACTGTAATGAATGTCTCAGGAGGACAGAGCTCAA 82
Qy 426 GlyArgAspThrLeuAsnThrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMet 445
Db 83 GCCCGAGACAGCGCTCAACTATTGTGCAGAGGCTTGGGAGGAATATATTGCGTATATG 142
Qy 446 GlnGluAlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsn 465
Db 143 CAAGAGCAAGAGTATGATCGCCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 202
Qy 466 GlyLysValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIle 485
Db 203 GGGAAAGTTAGCTCTGGTCTCGGATCGGATTCGACCCATTCCTGACACCGACATC 262
Qy 486 ProPheProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeulle 505
Db 263 CCCTTTCTGTGACGACGCTCTCAAGGAGTTGACATTCATTCGAGCTCAATGACTTGCA 322
Qy 506 CysIleIleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGly 525

```

```

Db 323 TCTGCCATTCTTCGATTACAGGGGATACGCGCTGTACACGCGACAGGCGCCGTGGA 382
Qy 526 GluGluAlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAsp 545
Db 383 GAAGAAGCTTCGTGTATATCTTGTATATGAAGAACAATCTCGGAACAACACAGAGGAAGAT 442
Qy 546 AlaLeuAsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeu 565
Db 443 GCTCTCAATCATCTCAACGCCATGATCAGTGATGATTAATAAGGATTAAATTTGGAGCTT 502
Qy 566 LeuLysProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArg 585
Db 503 CTCAAAATCAACAGCAGCGGTTCCCATATCTGCCAAAAACAACATGCTTTTGACATTAGCAGA 562
Qy 586 ValTrpHisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLys 605
Db 563 GCTTTCCATTGGCTCAAAATATCGAGATGCTACAGCGTTGCCAACATTAACAAACNNNN 622
Qy 606 SerLeuValMetArgThrValIleGluProValProLeu 618
Db 623 NGTTTGGTGANNNNACCGTCATTGATCTGTCACTTTA 661

RESULT 6
BX680641
LOCUS BX680641 RS Pinus pinaster cDNA clone RS46D06, mRNA sequence.
DEFINITION BX680641
ACCESSION BX680641
VERSION BX680641.1 GI:38015099
KEYWORDS EST.
SOURCE Pinus pinaster
ORGANISM Pinus pinaster
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE 1 (bases 1 to 599)
AUTHORS Frigerio, J. and Plomion, C.
TITLE Identification of water-deficit responsive genes in Maritime pine
(Pinus pinaster Ait.) using an EST approach
JOURNAL Unpublished (2002)
COMMENT Contact: Frigerio JM
Genetique et Amelioration 69
INRA
route d'Arcachon 33612 Cestas CEDEX France
Email: Frigerio@pierrot.inra.fr
Seq primer: T3.
FEATURES             Location/Qualifiers
     source
       1..599
         /organism="Pinus pinaster"
         /mol_type="mRNA"
         /db_xref="taxon:71647"
         /clone="RS46D06"
         /tissue_type="root"
         /dev_stage="6 weeks old seedling"
         /lab_host="SOLR"
         /clone_lib="RS"
         /note="Vector: Uni-ZAP XR; ecotype: Landes; The library
was made from the roots of 6 weeks old seedlings grown in
hydroponic conditions. A three weeks drought stress
treatment was applied by lowering the osmotic potential of
the nutrient solution to -0.45 MPa using 170 g/l of
polyethylene glycol as an osmoticum. A mixture of
genotypes were used. Oligo-dT primed cDNA was
directionally cloned into the EcoRI-XhoI lambda-ZAP vector
arms and mass-excised to form a pBluescript phagemid"
ORIGIN
Alignment Scores:
Pred. No.:          1,31e-87          Length:          599
Score:             853.00             Matches:         161
Percent Similarity: 87.94%             Conservative:    14
Best Local Similarity: 80.90%           Mismatches:      24
Query Match:       26.24%             Indels:          1

```

```
DB: 13 Gaps: 0
US-10-025-145A-65 (1-618) x BX680641 (1-599)
QY 399 LeuProGluThrMetLysGlyValTyrMetMetValTyrHisThrValAsnGluMetAla 418
Db 3 CTTCCCAATATATGAAGAAT-TACATGATGGTTTACACACCGTAATGAATGTCT 61
QY 419 ArgValAlaGluLysAlaGlnGlyArgAspThrLeuAsnTyrAlaArgGlnAlaTrpGlu 438
Db 62 GCGGAGGCGCAGAGGCTCAAGGCCGAGACACTCTCAACTATGCTCGACAGGCTGGGAG 121
QY 439 AlaCysPheAspSerTyrMetGlnGluAlaLysTrpIleAlaThrGlyTyrLeuProThr 458
Db 122 GATTATCTTGATTCGTATGATGCAAGCAAGCAAGTGGATAGCCAGCGTTATCTGCCAACG 181
QY 459 PheGluGluTyrLeuGluAsnGlyLysValSerSerAlaHisArgProCysAlaLeuGln 478
Db 182 TTCAGGATATCTGGAGAACGGGAAGTAGCTCTGGGATCGCTGTCGGCGTTGCAA 241
QY 479 ProIleLeuThrLeuAspIleProPheProAspHisIleLeuLysGluValAspPhePro 498
Db 242 CCCATGCTGACGATGACATCCCTCTTCTCTCACATCTCAAGGAAGTTGACTTCCCA 301
QY 499 SerLysLeuAsnAspIleCysIleLeuArgLeuArgGlyAspThrArgCysTyr 518
Db 302 TCCAACTCAATGACTTGGCATGTGCATACCTTCGATTACGAGTGATACCGATGTTAT 361
QY 519 LysAlaAspArgAlaArgGlyGluGluAlaSerSerIleSerCysTyrMetLysAspAsn 538
Db 362 CAGGAGCAGAGGCTCGTGGAGAGAAACCTCGTGATATCTGTTATATGAAGACAAC 421
QY 539 ProGlyLeuThrGluGluAspAlaLeuAsnHisIleAsnPheMetIleArgAspAlaIle 558
Db 422 CTTGGAGCAACAGAGGAAGATGCTCTTAATCATCTCAATCTCATGATCACTGCGTAAT 481
QY 559 ArgGluLeuAsnTrpGluLeuLysProAspAsnSerValProIleThrSerLysLys 578
Db 482 AAGAATAAATTTGGGAGCTCTCAAAACCCGACACACAGTGTGCCATTTCTTCCAGAA 541
QY 579 HisAlaPheAspIleSerArgValTrpHisGlyTyrArgTyrArgAspGlyTyr 597
Db 542 ATACTTTTGACATTACAGAGCTTTCCATTACGGTTACAAATACCGAGATGGCTAC 598

RESULT 7
BG526917
LOCUS
DEFINITION
  NXPV 057 D04 F NXPV (Nsf Xylem planings wood Vertical) Pinus taeda
  cDNA clone NXPV 057 D04 5' similar to Arabidopsis thaliana sequence
  At4g16730 limonene cyclase like protein see
  http://mips.gsf.de/proj/thal/db/index.html, mRNA sequence.
ACCESSION
  BG526917
VERSION
  BG526917.1 GI:13536796
KEYWORDS
  EST.
SOURCE
  Pinus taeda (loblolly pine)
  Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
  1 (bases 1 to 599)
  Sederoff, R.
  Molecular Basis of Wood Formation in the Pine Megagenome
  Unpublished (2000)
  Contact: Sederoff, Ron
  Forest Biotechnology
  North Carolina State University
  840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,
  NC 27695, USA
  Tel: 919 515 7800
  Fax: 919 515 7801
  Email: ron_sederoff@ncsu.edu, jerri_johnson@ncsu.edu
  Please see http://web.ahc.umn.edu/biodata/nsfpine/ for further
  information.
  Seq primer: T3.
```

## FEATURES

Source

## Location/Qualifiers

```
1..599
/organism="Pinus taeda"
/mol_type="mRNA"
/strain="Coastal plain loblolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="NXPV 057 D04"
/tissue_type="Xylem"
/cell_type="planings (secondary)"
/dev_stage="Transitional"
/lab_host="XL1-Blue"
/notes="Vector: Bluescript SK; Site 1: Eco RI; Site 2:
XhoI; The library is from early (spring) secondary wood,
taken from a ten year old tree in the transitional phase.
The tree is a kind gift of the Westvaco Corporation.
Secondary xylem was harvested from the tree by peeling
back the bark and primary xylem and then removing the
underlying tissue with a block plane. NOTE: The sequences
contain a 'cDNA adapter' between the EcoRI site and the
start of the EST. The adapter sequence is
'AATTCGGCAGG'."
```

## ORIGIN

## Alignment Scores:

```
Pred. No.: 1,04e-81 Length: 599
Score: 802.00 Matches: 150
Percent Similarity: 83.92% Conservative: 17
Best Local Similarity: 75.38% Mismatches: 32
Query Match: 24.67% Indels: 0
DB: 12 Gaps: 0
```

US-10-025-145A-65 (1-618) x BG526917 (1-599)

```
QY 395 AlaMetGluCysLeuProGluTyrMetLysGlyValTyrMetMetValTyrHisThrVal 414
Db 3 GCGACAGAGTGCCTCCAGAAATATATGAAGGAGTTTACATGATAGTTTACAACACTGTA 62
QY 415 AsnGluMetAlaArgValAlaGluLysAlaGlnGlyArgAspThrLeuAsnTyrAlaArg 434
Db 63 AATGAATGCTCAGAGGCGACACAGGCTCAAGGCCGAGACACGCTCACTATTGTCGA 122
QY 435 GlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGluAlaLysTrpIleAlaThrGly 454
Db 123 CAGGCTGGGAGGAATATATTGATGCTATATGCAAGCAAGCAAGTGCAGTCGTCAGTGT 182
QY 455 TyrLeuProThrPheGluGluTyrLeuGluAsnGlyLysValSerSerAlaHisArgPro 474
Db 183 GAGGTGCCAACATTTGAGGAGTACTATGAGAACGGGAAGTTAGCTCTGTCATCGCGTA 242
QY 475 CysAlaLeuGlnProIleLeuThrLeuAspIleProPheProAspHisIleLeuLysGlu 494
Db 243 TCGGCATTGCAACCCATTTTGACGACGACATCCCCCTTCTGAGCAGCTCTCAAGAA 302
QY 495 ValAspPheProSerLysLeuAsnAspLeuIleCysIleIleLeuArgLeuArgGlyAsp 514
Db 303 GTTGACATTCATCGCAGCTCAATGACTTGGCATCTGCCATCTCTTCGATTACGAGGGAT 362
QY 515 ThrArgCysTyrLysAlaAspArgAlaArgGlyGluGluAlaSerSerIleSerCysTyr 534
Db 363 ACAGCTGCTACCCNNGCGACAGGCCCGTGGAGGAAGCTTCGTGTATCTGTAT 422
QY 535 MetLysAspAsnProGlyLeuThrGluGluAspAlaLeuAsnHisIleAsnPheMetIle 554
Db 423 ATGAAAGCAATCTCTNNAAACAAGAGAGATGCTCTCAATCATCTCAAGCCATGATC 482
QY 555 ArgAspAlaIleArgGluLeuAsnTrpGluLeuLysProAspAsnSerValProIle 574
Db 483 AGTGATGTTATTANANNNTTAAATTGGGAGCTTCTCAAAACCAACAGCAGCTTCCCAT 542
QY 575 ThrSerLysLysHisAlaPheAspIleSerArgValTrpHisGlyTyrArgTyr 593
Db 543 TCTGCCAAANNCAATGCTTTTGACATTANCNNNNNNTCCNNTGTCGTACAAATAT 599
```

QY	16	SerSerSerHisGluLeuIleValAlaLeuArgThrlleProThrLeuGluIleCysArg	35
LOCUS	80	GGTTTTCATGAGCTGAAAGCTATCATAGTACAGTCCCAAAATCTTGAATATGCAGG	139
DEFINITION	36	ProGlyLysSerValAlaHisSerIleAsnMet---CysLeuThrSerValAlaSerThr	54
ACCESSION	140	CGAGGGAAATCCATAGCACCTTCATGAGCATGAGTTCGACCACCTCGGTTCCTAAAG	199
VERSION	55	AspSerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIle	74
KEYWORDS	200	GATGGGTACCAAGACGATAGCTGGTTCATCATTTCCAACTTTTGGACGATGATTCATTA	259
SOURCE	75	GlnSerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeu	94
ORGANISM	260	GCCTCTCTC---TCCACTTCCTATGAGCACCTTCTACCGTACGCGCTGATAAACTT	316
REFERENCE	95	IleGlyGluValLysAspIleMetPheAsnPhelysSerLeuGluAspGlyGly	112
AUTHORS	317	ATAGGGGAAGTAAATAATATC---TTCGATTAAATGTAGTGGAGGATGAGTATTCAC	373
TITLE	113	-----AsnAspLeuGlnArgLeuLeuValAspValGluArgLeuGly	129
JOURNAL	374	AGTCCCTCAGTGACCTCCATCCGCTCTGGATGGTGGATAGCGGTGAACGGTTGGGA	433
COMMENT	130	IleAspArgHisPheLysLysGluLeuIleLysThrAlaLeuAspTyrValAsnSerTyrTrp	149
	434	ATCGATAGGCATTTCAAAGACGAGATAAATCTCTCTCGATCATGTTTACAGTATTGG	493
	150	AsnGluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAla	169
	494	ACCGAAAAGGCAATTTGACGGCGAGAGAAAGTGTGTGACTGATCTCAACTCAACTGCC	553
	170	LeuGlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnVal	189
	554	TTGGCCCTTCGAATCTCCGACTACAGGATACAGTGTCTTCAATGTTCTGGATCAC	613
	190	PhelysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnLeuGlyGluIle	209
	614	TTCAAAAACGAGAAGGGCGAGTTTACTTGTCTGCTGCC---ATTCAAAACAGAGAGATA	670
	210	ArgGlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMet	229
	671	AGAGATGTCCTCAATTTATTTGGGCACTCTCTCATTTGCCCTTCTGGCAAAAATTATG	730
	230	AspGluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAla	249
	731	GAGCGCGTGAAATCTTCTCTCAATGATATTAAAGATGCCCTACAAAAGATTCGCC	790
	250	SerSerIle 252	
	791	TCCGGTCTT 799	
FEATURES	CF663845	616 bp mRNA linear EST 07-OCT-2003	
source	RTCN1_5_B08_g1_A029	Root control Pinus taeda cDNA clone	
	CF663845	RTCN1_5_B08_A029 5', mRNA sequence.	
	CF663845	GI:37561088	
	CF663845.1	GI:37561088	
	Pinus taeda (loblolly pine)		
	Pinus taeda		
	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.		
	1 (bases 1 to 616)		
	Pratt,L., Cordomier-Pratt,M.-M., Lorenz,W.W., Zimmermann,C. and		
	Dean,J.F.D.		
	An EST database from untreated loblolly pine (Pinus taeda) roots		
	Unpublished (2003)		
	Other ESTs: RTCN1_5_B08_b1_A029		
	Contact: Cordomier-Pratt MM		
	Laboratory for Genomics and Bioinformatics		
	The University of Georgia, Department of Plant Biology		
	Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA		
	Tel: 706 542 1860		
	Fax: 706 583 0210		
	Email: mmpzatt@uga.edu		
	RNA prepared and library constructed by W. Walter Lorenz (School of		
	Forest Resources, University of Georgia); plant material prepared		
	by Craig Zimmermann (School of Forest Resources, University of		
	Georgia) using rooted cuttings provided by the Forest Biology		
	Research Cooperative (FBRC) and the CCIONES project a the		
	University of Florida; sequencing done in the Laboratory for		
	Genomics and Bioinformatics, University of Georgia. Sequence ends		
	have been trimmed to exclude vector and regions below fired quality		
	16. Three-prime sequences are presented as their reverse complement		
	and have been trimmed to exclude polyA.		
	Seq primer: JENREV (CAGGAACAGCTATGACC).		
	Location/Qualifiers		
	1. .804		
	/organism="Pinus taeda"		
	/mol_type="mRNA"		
	/strain="3 CCIONES"		
	/db_xref="taxon:3352"		
	/clone="RTCN1_22_C05_A029"		
	/lab_host="DH10B-T1 phage-resistant E. coli"		
	/clone_lib="Root control"		
	/note="organ: root; Vector: pSL1180; Site_1: EcoRI;		
	from the roots of 1-year-old loblolly pine (Pinus taeda)		
	cuttings that were rooted and then planted in washed sand.		
	Just before harvesting roots for RNA isolation, the rooted		
	cuttings were maintained for 27 days (April 2003) under		
	ambient conditions in a local greenhouse. They were kept		
	on a weekly regimen of 0.5x nutrient-complete Hoagland's		
	solution and supplemented with additional water sufficient		
	to maintain a 15% soil moisture content. Double-stranded		
	cDNA was cloned unidirectionally into pSL1180. Inserts can		
	be excised with EcoRI (5' end) and XhoI (3' end)."		
ALIGNMENT	CF666338	804 bp mRNA linear EST 07-OCT-2003	
source	RTCN1_22_C05_g1_A029	Root control Pinus taeda cDNA clone	
	CF666338	RTCN1_22_C05_A029 5', mRNA sequence.	
	CF666338.1		

The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
RNA prepared and library constructed by W. Walter Lorenz (School of Forest Resources, University of Georgia); plant material prepared by Craig Zimmermann (School of Forest Resources, University of Georgia) using rooted cuttings provided by the Forest Biology Research Cooperative (FBRC) and the CCLONES project at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACAGCTATGACC).

# FEATURES

Location/Qualifiers  
1. .616  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="3 CCLONES"  
/db\_xref="taxon:3352"  
/clone="R1CMT1 5 B08 A029"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="Root control"  
/note="Organ: root; Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from the roots of 1-year-old loblolly pine (Pinus taeda) cuttings that were rooted and then planted in washed sand. Just before harvesting roots for RNA isolation, the rooted cuttings were maintained for 27 days (April 2003) under ambient conditions in a local greenhouse. They were kept on a weekly regimen of 0.5x nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain a 15% soil moisture content. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts can be excised with EcoRI (5' end) and XhoI (3' end)."

# ORIGIN

Alignment Scores:  
Pred. No.: 7,81e-80 Length: 616  
Score: 786.00 Matches: 143  
Percent Similarity: 83.08% Conservative: 24  
Best Local Similarity: 71.14% Mismatches: 34  
Query Match: 24.18% Indels: 0  
DB: 14 Gaps: 0

US-10-025-145A-65 (1-618) x CF663845 (1-616)

QY 286 LysAsnLysAsnAlaAlaGluLysLeuLeuGluLeuAlaLysLeuGluPheAsnIlePhe 305  
DB 12 AAGACGACGACCGAGAGAGCTTTTGAAGCTTGCAGAAATTTGGAGTTCAATATCTTT 71  
QY 306 HisSerLeuGluGluArgGluLeuLysHisValSerArgTrpTrpLysAspSerGlySer 325  
DB 72 CACTCTTCACGACAAAGAGAGTTAAACACAGCTGCCAGATGGTGGGAAGATTTCGGGTTC 131  
QY 326 ProGluMetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIle 345  
DB 132 TCTCAACTGCATTCACTCGCGCATCGTCAGTGCAGTTCACATTTGGCCCTCTCGATT 191  
QY 346 AlaPheGluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIle 365  
DB 192 GCCACTGAGCCCAACATTCAGCATTCAGTTGGCTTGCAGAAACGCTGTTAICTTGA 251  
QY 366 ThrValLeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThr 385  
DB 252 ATAGTTCTGGACGACATCTATGACATCTTCGGAACAATGGAGAGCTCGAATCTTCACA 311  
QY 386 AlaThrIleLysArgTrpAspProSerAlaMetGluCysLeuProGluTyrMetLysGly 405  
DB 312 GCCGCAATTAAAGATGGATCCCGTCGCGGAGGAGTTCCTTCAGAAATATATGAAGGC 371

QY 406 ValTyrMetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGln 425  
DB 372 ATATATATGGTCTTTTACGATTGCTTAATCAATGGCTCGAGCGCGGAGAGATCAA 431  
QY 426 GlyArgAspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMet 445  
DB 432 GCCCGAGACACCTCACCTACCTCGAATACGTGGGAAGCCGTATTGTGATGGTTCTG 491  
QY 446 GlnGluAlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsn 465  
DB 492 GAAGAGACAAATGGATGTCAGCGGTATATATACCCACGTTTGAGGAGTATTGGAGAAC 551  
QY 466 GlyLysValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIle 485  
DB 552 GCGMAGGTGAGTTTGGGTATCGAGCAGCCACACTGCAACCCATCTCAGTTGGATATT 611  
QY 486 Pro 486  
DB 612 CCC 614

# RESULT 10

CF476978 637 bp mRNA linear EST 08-SEP-2003  
RTW3\_5\_A06\_b1\_A022 Well1-watered loblolly pine roots WM3 Pinus  
taeda cDNA clone RTW3\_5\_A06\_A022 3', mRNA sequence.  
CF476978  
CF476978.1 GI:34505847  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Pinus taeda (loblolly pine)  
Pinus taeda  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
1 (bases 1 to 637)  
Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J.,  
Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and  
Neale, D.  
An EST database from well-watered loblolly pine (Pinus taeda) roots  
Unpublished (2003)  
Other ESTs: RTW3\_5\_A06-gl\_A022  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
RNA prepared and library constructed by W. Walter Lorenz, School of  
Forestry, University of Georgia; plant material prepared at the  
University of Florida; sequencing done in the Laboratory for  
Genomics and Bioinformatics, University of Georgia. Sequence ends  
have been trimmed to exclude vector and regions below phred quality  
16. Three-prime sequences are presented as their reverse complement  
and have been trimmed to exclude polyA.  
Seq primer: M13-21 (TGTAACACGACGCCAGT)  
POLYA=Yes.

# FEATURES

Location/Qualifiers  
1. .637  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="CCLONES"  
/db\_xref="taxon:3352"  
/clone="RTW3\_5\_A06\_A022"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="Well1-watered loblolly pine roots WM3"  
/note="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The  
library was prepared from polyA+ RNA from loblolly pine  
(Pinus taeda) roots watered to pot capacity every other  
day. Pre-dawn water potential remained -0.3 MPa +/-0.1.  
Roots were harvested for RNA isolation. Double-stranded  
cDNA was cloned unidirectionally into pSL1180. Inserts  
excised with EcoRI (5' end) and XhoI (3' end)."

# ORIGIN

## Alignment Scores:

Pred. No.: 2,92e-77 Length: 637  
 Score: 764.00 Matches: 142  
 Percent Similarity: 86.41% Conservativity: 17  
 Best Local Similarity: 77.17% Mismatches: 25  
 Query Match: 23.50% Indels: 0  
 DB: 14 Gaps: 0

US-10-025-145A-65 (1-618) x CF476978 (1-637)

QY 435 GlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGluAlaLysTrpIleAlaThrGly 454  
 Db 3 CAGGCTTGGAGGAATATATGCTATATGCGATATGCGAAGCAAGTGGATCCCGAGTGT 62  
 QY 455 TyrLeuProThrPheGluGluTyrLeuGluAsnGlyLysValSerSerAlaHisArgPro 474  
 Db 63 GAGGTGCCAACATTTGAGGAGTACTACGAGAACGGGAAATAGTCTGGTCATCGCGTA 122  
 QY 475 CysAlaLeuGlnProIleLeuThrLeuAspIleProPheProAspHisIleLeuLysGlu 494  
 Db 123 TCGGATTCGAACCATTTCTGACGCCGACATCCCTTTCTCTGAGCAGCTCCTCAAGGAA 182  
 QY 495 ValAspPheProSerLysLeuAsnAspLeulleCysIleIleLeuArgLeuArgGlyAsp 514  
 Db 183 GCTGACATTCATGCAAGCTCAATGACTTGGCATCTGCCATCTTCGATTCGATTACGAGAGAT 242  
 QY 515 ThrArgCysTyrLysAlaAspArgAlaArgGlyGluGluAlaSerSerIleSerCysTyr 534  
 Db 243 ACGGCTGCTACAGCGCGGACAGGCGCGTGGAGAAAGCTTCGTGTATATCTTGTAT 302  
 QY 535 MetLysAspAsnProGlyLeuThrGluGluAspAlaLeuAsnHisIleAsnPheMetIle 554  
 Db 303 ATGAAGACATATCTGGAGCAACGGAAGAGATGCTCTCAATCATATCAACGCCATGATC 362  
 QY 555 ArgAspAlaIleArgGluLeuAsnTrpGluLeuLysProAspAsnSerValProIle 574  
 Db 363 AGTGATGTAATTAAGGATTAATTTGGAGCTTCTCAACCAACAGCAGCGTCCCAT 422  
 QY 575 ThrSerLysLysHisAlaPheAspIleSerArgValTrpHisHisGlyTyrArgTyrArg 594  
 Db 423 TCTCCCAAAAACATGCTTTTGACATAGCAGAGCTTTCATATATGCTTACAAATATCGA 482  
 QY 595 AspGlyTyrSerPheAlaAsnValCluThrLysSerLeuValMetArgThrValIleGlu 614  
 Db 483 GATGGCTACAGCTTCCGACATTCGAACAAAGAGTTTGGTGAACCGCAACCGTCATTGAT 542  
 QY 615 ProValProLeu 618  
 Db 543 GCTGTGACTTTA 554

RESULT 11  
 AL750951  
 LOCUS AL750951 RS pinus pinaster cDNA clone RS02D01 linear EST 20-JUN-2002  
 DEFINITION SYNTHASE, mRNA sequence.

ACCESSION AL750951  
 VERSION AL750951.1  
 KEYWORDS EST.  
 SOURCE Pinus pinaster

## ORGANISM

Pinus pinaster  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

1 (bases 1 to 481)  
 Frigerio, J. and Plomion, C.

Identification of water-deficit responsive genes in Maritime pine  
 (Pinus pinaster Ait.) using an EST approach

Unpublished (2002)  
 Contact: Frigerio JM

Genetique et Amelioration 69  
 INRA

route d'Arcachon 33612 Cestas CEDEX France  
 Email: Frigerio@pierrot.inra.fr

Seq primer: T3.

## FEATURES

## source

## Location/Qualifiers

1. .481  
 /organism="Pinus pinaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:71647"  
 /clone="RS02D01"  
 /tissue\_type="root"  
 /dev\_stage="6 weeks old seedling"  
 /lab\_host="SOLP"  
 /clone\_lib="RS"

/note="Vector: Uni-ZAP XR; ecotype: Landes; The library was made from the roots of 6 weeks old seedlings grown in hydroponic conditions. A three weeks drought stress treatment was applied by lowering the osmotic potential of the nutrient solution to -0.45 MPa using 170 g/l of polyethylene glycol as an osmoticum. A mixture of genotypes were used. Oligo-dT primed cDNA was directionally cloned into the EcoRI-XhoI lambda-ZAP vector arms and mass-excised to form a pBluescript phagemid"

## ORIGIN

## Alignment Scores:

Pred. No.: 6,76e-74 Length: 481  
 Score: 733.00 Matches: 134  
 Percent Similarity: 90.62% Conservativity: 11  
 Best Local Similarity: 83.75% Mismatches: 15  
 Query Match: 22.55% Indels: 0  
 DB: 9 Gaps: 0

US-10-025-145A-65 (1-618) x AL750951 (1-481)

QY 335 HisValGluTyrTyrAlaLeuAlaSerCysIleAlaPheGluProGlnHisSerGlyPhe 354  
 Db 2 CACGTGGAGTACTACACTTTCGCTTCCTGCAATCGGTTTGAGCCCCAACATCTCGATTC 61  
 QY 355 ArgLeuGlyPheThrLysMetSerHisLeuIleThrValLeuAspMetTyrAspVal 374  
 Db 62 AGACTCGGCTTTGGGAAAGCGTGCATATTATCACTGTTCGACGATATGTACGACCTC 121  
 QY 375 PheGlyThrValAspGluLeuGluPheThrAlaThrIleLysArgTrpAspProSer 394  
 Db 122 TTCGGAACAGTTGTAAGAGCTCAAACTGTTCACGCCCAATTAAAGATGGGATCCGTC 181  
 QY 395 AlaMetGluCysLeuProGluTyrMetLysGlyValTyrMetValTyrHisThrVal 414  
 Db 182 GCCACAGATTGCTTCACATATATGAAGGAATTTACATGATGGTTTACACACCGTA 241  
 QY 415 AsnGluMetAlaArgValAlaGluLysAlaGlnGlyArgAspThrLeuAsnTyrAlaArg 434  
 Db 242 AATGAAATGCTGGGAGGCGACAGAGGCTCAAGGCCGAGACACTCTCAACTATGCTGA 301  
 QY 435 GlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGluAlaLysTrpIleAlaThrGly 454  
 Db 302 CAGGCTTGGGAGGAGTATCTTGATTCGTATATGCAAGAAGCAAGTAGTGATGCCCGGT 361  
 QY 455 TyrLeuProThrPheGluGluTyrLeuGluAsnGlyLysValSerSerAlaHisArgPro 474  
 Db 362 TATCTGCAACGTTCCGAGGAATACTTGGAGAACCGGAAAGTAGTCTTGGGCATCGGTG 421  
 QY 475 CysAlaLeuGlnProIleLeuThrLeuAspIleProPheProAspHisIleLeuLysGlu 494  
 Db 422 TCGGCGTTGCAACCCATGCTGAGCATGACATCCCTTTCTCTCATCATCTCAAGGAA 481

RESULT 12  
 CF477562/c

LOCUS CF477562 740 bp mRNA linear EST 08-SEP-2003

DEFINITION RTWW3\_8\_G10\_g1\_A022 Well-watered loblolly pine roots WW3 Pinus taeda cDNA clone RTWW3\_8\_G10\_A022 5', mRNA sequence.

ACCESSION CF477562  
 VERSION CF477562.1  
 KEYWORDS EST.  
 SOURCE Pinus taeda (loblolly pine)

ORGANISM Pinus taeda

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

1 (bases 1 to 740)

Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W. M., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J., and Neale, D.

An EST database from well-watered loblolly pine (Pinus taeda) roots Unpublished (2003)

Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.

Seq primer: JENREV (CAGGAACAGCTATGACC).

#### FEATURES

Source

1..740

/organism="Pinus taeda"

/mol\_type="mRNA"

/strain="CCIONES"

/db\_xref="taxon:3352"

/clone="RTW3 8 G10 A022"

/lab\_host="DH10B-T1 phage-resistant E. coli"

/clone\_lib="well-watered loblolly pine roots W3"

/notes="Vector: pSL180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/- 0.1.

Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

#### ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

2,14e-72

723.00

84.54%

66.49%

22.24%

14

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

740

129

35

30

0

0

US-10-025-145A-65 (1-618) x CF477562 (1-740)

425 GlnGlyArgAspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyr 444

739 CAAGGCTGGGATACAGTCAGTCATGCTCGAAATCTTGGGAGGCTTATTTGGTGCTTAT 680

445 MetGlnGluAlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGlu 464

679 ATACAGAGCCCAAGTGGATTTTCAGTGGTATCTTCCACGTTCCGAGTACCTCGAG 620

465 AsnGlyLysValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAsp 484

619 AATGGGAAGTCCAGTTCGGCTCTCGCATTAACACCGCTCGAACCCACGTCGACTTGGGG 560

485 IleProPheProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeu 504

559 TTTCTCTTCGCGCTCGAATCTCGAGAAATTCGACTTTCATCGAATTCGAATTCGAATTTG 500

505 IleCysIleIleLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArg 524

499 ATATGTGCCATCTCTCGACTGAAGGTGACACTCAATGCTACAAGGCTGACAGGCGCGT 440

525 GlyGluGluAlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGlu 544

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

Db 439 GGAGAGAGAGCTTCGCCGCTATCGTGTATTATGAAGACCATCTCGGAATAACAGAGAA 380

QY 545 AspAlaLeuAsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGlu 564

Db 379 GATGCTGCAATCAAGTCAATGCTATGTCGTAATACTTAACCAAGAACTGAATGGGAG 320

QY 565 LeuLeuLysProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSer 584

Db 319 TTACTTAGACCCGACAGCGGTGTTCCCATCTCTTACAGAGAGTTCGCTTTTGCACATTTGC 260

QY 585 ArgValTrpHisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThr 604

Db 259 AGAGTTTTCATTACGTTTACAAATACAGAGATGGCTTCAGTGTGCCAGTATTGAATA 200

QY 605 LysSerLeuValMetArgThrValIleGluProValProLeu 618

Db 199 AAGATTGTTGTAACAGAACCGTGGTTGAACATGTCCTTTG 158

RESULT 13

LOCUS

DEFINITION

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

BO698077

#### ORIGIN

Alignment Scores:

Pred. No.:

5.46e-72

Length:

516

#### FEATURES

source

1..516

/organism="Pinus taeda"

/mol\_type="mRNA"

/strain="Coastal plain loblolly pine from North Carolina"

/db\_xref="taxon:3352"

/clone="NXPV 064 C05"

/tissue\_type="Xylem"

/cell\_type="Planings (secondary)"

/dev\_stage="Transitional"

/lab\_host="XLI-Blue"

/clone\_lib="NXPV (Nsf Xylem Planings wood Vertical)"

/note="Vector: Bluescript SK; Site 1: Eco RI; Site 2: XhoI; The library is from early (spring) secondary wood, taken from a ten year old tree in the transitional phase. The tree is a kind gift of the Westvaco Corporation. Secondary xylem was harvested from the tree by peeling back the bark and primary xylem and then removing the underlying tissue with a block plane. NOTE: The sequences contain a 'cDNA adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'AATTCGGCAGAG'."

Seq primer: T3.

Location/Qualifiers

```

/tissue_type="root"
/dev_stage="6 weeks old seedling"
/lab_host="SOUR"
/clone_lib="RN"
/notes="Vector: Uni-ZAP XR; ecotype: Landes; The library
was made from the toots of 6 weeks old seedlings grown in
hydroponic conditions. A mixture of genotypes were used.
Oligo-dT primed cDNA was directionally cloned into the
EcoRI-XhoI lambda-ZAP vector arms and mass-excised to form
a pBluescript phagemid"

ORIGIN

Alignment Scores:
Pred. No.: 1 59e-71 Length: 517
Score: 713.00 Matches: 136
Percent Similarity: 86.63% Conservative: 13
Best Local Similarity: 79.07% Mismatches: 23
Query Match: 21.93% Indels: 1
DB: 13 Gaps: 0

US-10-025-145A-65 (1-618) x BX677624 (1-517)

QY 394 SerAlaMetGluCysLeuProGluTyrMetLysGlyValTyrMetMetValTyrHisThr 413
Db 1 TCGCGCAGACAGTGCCTTCAGAAATATATGAAAGGAGT-TACATGATAGTITACACACT 59
QY 414 ValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArgAspThrLeuAsnTyrAla 433
Db 60 ATAAATGAAATGCTCAGGAGGAGCAACAGGCTCAAGGCGGAGACACGCTCAACTATTGT 119
QY 434 ArgGlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGluAlaLysTrpIleAlaThr 453
Db 120 CGACAGCTTGGAGGAATATATTGATGGTATATGCAAGACCAAGTGGATCGGCAGT 179
QY 454 GlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLysValSerSerAlaHisArg 473
Db 180 GGTGAGGTGCCAACATTTGAGGAGTACTATGAGACGGGAAAGTTAGCTTGGTCATCGC 239
QY 474 ProCysAlaLeuGlnProIleLeuThrLeuAspIleProPheProAspHisIleLeuLys 493
Db 240 GTATCGGCATTGCACCCATTCTGACGACCGCATCCCTTCTCTGAGCACGCTCTCAAG 299
QY 494 GluValAspPheProSerLysLeuAsnAspLeuIleCysIleIleLeuArgLeuArgGly 513
Db 300 GAAGTGGACATTCATCGAAGCTCAATGACTTGGCATCTGGCCATCTTCGATTACGAGG 359
QY 514 AspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGluAlaSerSerIleSerCys 533
Db 360 GACACTCGCTGTACCGAGCGGACAGGCGCGTGGAGAGAAGTTCGGGTATATCTTCT 419
QY 534 TyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeuAsnHisIleAsnPheMet 553
Db 420 TATATGAAGACAAATCTCGAACACAGAGGAAGATGCTCTCAATCATATCAACCCCATG 479
QY 554 IleArgAspAlaIleArgGluLeuAsnTrpGluLeu 565
Db 480 ATCAGTGATGTAATTAAGATTAATTTGGGAGCTT 515

RESULT 15
AW287756 539 bp mRNA linear EST 07-JAN-2000
LOCUS Sitka spruce drill-wounded bark Picea sitchensis cDNA clone
DEFINITION 25-1-3 '5' similar to mono-terpene synthase, mRNA sequence.
ACCESSION AW287756
VERSION AW287756.1 GI:6681768
KEYWORDS EST.
SOURCE Picea sitchensis (Sitka spruce)
ORGANISM Picea sitchensis
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
REFERENCE 1 (bases 1 to 539)
AUTHORS Wang,S.X., Hunter,W. and Plant,A.L.
TITLE Isolation of terpene synthase gene-specific probes from Sitka

Score: 717.00 Matches: 133
Percent Similarity: 88.17% Conservative: 16
Best Local Similarity: 78.70% Mismatches: 20
Query Match: 22.05% Indels: 0
DB: 13 Gaps: 0

US-10-025-145A-65 (1-618) x BQ698077 (1-516)

QY 442 AspSerTyrMetGlnGluAlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGlu 461
Db 8 GATCGGTATATGCAAGAAGCAAGTGGATCGCCAGTGGTGAAGTGCACCAACTTTGAGAG 67
QY 462 TyrLeuGluAsnGlyLysValSerSerAlaHisArgProCysAlaLeuGlnProIleLeu 481
Db 68 TACTATGAGAAACGGGAAAGTTAGCTCTGTCATCGCGTATCGGCATTGCAACCCATTGT 127
QY 482 ThrLeuAspIleProPheProAspHisIleLeuLysGluValAspPheProSerLysLeu 501
Db 128 ACGACCCGACATCCCTTCTTGACGACGCTCTCAAGGAAGTGCATTCCTCGACGCTC 187
QY 502 AsnAspLeuIleCysIleIleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAsp 521
Db 188 AATGACTTGGCATCTGCCATCTCTTCGATTACGAGGGATACGCGCTCTACCGCGGAC 247
QY 522 ArgAlaArgGlyGluGluAlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeu 541
Db 248 AGGCCCGGTGGAGAAGAGTCTCGTATATCTTGTATATATGAAAGACAATCCTCGGAACA 307
QY 542 ThrGluGluAspAlaLeuAsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeu 561
Db 308 ACAGAGAAAGATGCTCTCAATCATCTCAAGCCATGATCATGATGTAATAAAGGATTA 367
QY 562 AsnTrpGluLeuLysProAspAsnSerValProIleThrSerLysLysHisAlaPhe 581
Db 368 AATTGGAGCTTCTCAAAACCAACAGCAGCGTCCCATATCTGCCAAAAAATGCTTTT 427
QY 582 AspIleSerArgValTrpHisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsn 601
Db 428 GACATTAGCAGAGCTTTCCCATGTTGGCTACAAATATCGAGATGGCTACAGCGTTGCCAAC 487
QY 602 ValGluThrLysSerLeuValMetArg 610
Db 488 ATTGAACAAGAGTTTGGTGAAGAGA 514

RESULT 14
BX677624 517 bp mRNA linear EST 28-OCT-2003
LOCUS RN Pinus pinaster cDNA clone RN42B08, mRNA sequence.
DEFINITION BX677624
ACCESSION BX677624
VERSION BX677624.1 GI:38011576
KEYWORDS EST.
SOURCE Pinus pinaster
ORGANISM Pinus pinaster
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE 1 (bases 1 to 517)
AUTHORS Prigiero,J. and Piomoni,C.
TITLE Identification of water-deficit responsive genes in Maritime pine
(Juniperus pinaster Ait.) using an EST approach
JOURNAL Unpublished (2002)
COMMENT Contact: Prigiero JM
Genetique et Amelioration 69
INRA
route d'Arcachon 33612 Cestas CEDEX France
Email: Prigiero@pierrotton.inra.fr
Email: Prigiero@pierrotton.inra.fr
Seq primer: T3
Location/Qualifiers
i. .517
/organism="Pinus pinaster"
/mol_type="mRNA"
/db_xref="taxon:71647"
/clone="RN42B08"

FEATURES
source

```

spruce and induction of gene expression by simulated white pine weevil damage

# JOURNAL COMMENT

Unpublished (2000)  
Contact: Wang SX  
Department of Biological Sciences  
Simon Fraser University  
8888 University Drive, Burnaby, BC V5A1S6, Canada  
Tel: 604 584 8870  
Fax: 604 584 8873  
Email: sxwang@bigfoot.com  
Clone sequence of a RT-PCR product from the mRNA of drill-wounded bark tissue

## PCR PRIMERS

FORWARD: Mult-F10  
BACKWARD: Mult-R18  
Insert Length: 539 Std Error: 0.00  
Seq primer: M13 Reverse and M13 Forward.  
High quality sequence stop: 539.

## FEATURES

source

1. 539  
/organism="Picea sitchensis"  
/mol\_type="mRNA"  
/db\_xref="taxon:3332"  
/clone="25-1-3"  
/tissue\_type="bark"  
/clone\_lib="Sitka spruce drill-wounded bark"  
/note="mRNA isolated from drill-wounded bark tissues;  
RT-PCR product"

## ORIGIN

### Alignment Scores:

Pred. No.:	9.35e-70	Length:	539
Score:	698.00	Matches:	126
Percent Similarity:	84.09%	Conservative:	22
Best Local Similarity:	71.59%	Mismatches:	28
Query Match:	21.47%	Indels:	0
DB:	10	Gaps:	0

US-10-025-145A-65 (1-618) x AW287756 (1-539)

QY	368	LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluPheThrAlaThr	387
Db	3	CTCGACGACCTGTACGACACATTCGGAACAATCGAAGAAATCGAATCTTTCACAGAAGCA	62
QY	388	IleLysArgTyrAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr	407
Db	63	GTCCAGGAGATGGGATCGTCGAGACAGAGAGCCCTCCAGACTATATGAAAGGAGTGATC	122
QY	408	MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg	427
Db	123	ATGGTACTCTACGAAGCCCTAACTGAANNGCTCAAGAGCGGAGAAAACACAGGCCGA	182
QY	428	AspThrLeuAsnTyrAlaArgGlnAlaTyrGluAlaCysPheAspSerTyrMetGlnGlu	447
Db	183	GACACGCTCAACTATGCTCGAAAGGCTTGGAGATTTATCTTGATTCGTATATTCAGAA	242
QY	448	AlaLysTyrIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys	467
Db	243	GCAAAGTGGATCGCCAGTGGTATCTGCCAACATTTCCAGGAGTACTTTGAGACGGGAA	302
QY	468	ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe	487
Db	303	ATTAGCTCTGTTATCCGCGAGCGGCATTGACACCCATCTCACATTTGGAGTACCGCTT	362
QY	488	ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle	507
Db	363	CCTGAATACATCTTGAAGGGAATTGATTTCCATCGAGATTCAATGATTTGGCATCTTCC	422
QY	508	IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu	527
Db	423	TTCCCTTCGACTAAGAGGTGACACACGCTGCTACAGCGGATAGGCCCCGTGGAGAGAA	482
QY	528	AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGlu	543

Db 483 GCTTCGTGCATATCTTGCTACATGAAGGACCACCCACTGCTACATGAA 530

Search completed: July 8, 2004, 20:31:50  
Job time : 4334 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 8, 2004, 16:55:46 ; Search time 786 Seconds  
(without alignments)  
3832.137 Million cell updates/sec

Title: US-10-025-145A-65  
Perfect score: 3251  
Sequence: 1 MALLSITPLVRSCLSSSHE.....FANVETKSLVMRTVIEPVPL 618

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3183909 seqs, 2436941669 residues

Total number of hits satisfying chosen parameters: 6367818

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US10025145/runat 07072004 160150 27970/app query.fasta\_1.775  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct.THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USERS=US10025145 @CGN 1 1 511 @runat 07072004 160150 27970  
-NCPU=6 -ICPU=3 -NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

1	3251	100.0	2013	15	US-10-025-145A-64	Sequence 64, Appl
2	2453	75.5	2018	9	US-09-887-586A-19	Sequence 19, Appl
3	2453	75.5	2018	9	US-09-903-012-19	Sequence 19, Appl
4	2453	75.5	2018	10	US-09-900-797-19	Sequence 19, Appl
5	2453	75.5	2018	13	US-09-893-820-19	Sequence 21, Appl
6	2453	75.5	2018	14	US-10-041-007-21	Sequence 21, Appl
7	2453	75.5	2018	15	US-10-025-145A-3	Sequence 3, Appl
8	2264	69.6	2186	15	US-10-025-145A-66	Sequence 66, Appl
9	2200.5	67.7	2196	9	US-09-887-586A-29	Sequence 29, Appl
10	2200.5	67.7	2196	9	US-09-903-012-29	Sequence 29, Appl
11	2200.5	67.7	2196	10	US-09-900-797-29	Sequence 29, Appl
12	2200.5	67.7	2196	13	US-09-893-820-29	Sequence 29, Appl
13	2200.5	67.7	2196	14	US-10-041-007-25	Sequence 25, Appl
14	2200.5	67.7	2196	15	US-10-025-145A-1	Sequence 1, Appl
15	2197.5	67.6	2205	15	US-10-025-145A-31	Sequence 31, Appl
16	2137	65.7	1890	15	US-10-025-145A-77	Sequence 77, Appl
17	2084.5	64.1	2429	14	US-10-041-007-27	Sequence 27, Appl
18	2084.5	64.1	2429	15	US-10-025-145A-68	Sequence 68, Appl
19	2031.5	62.5	2089	9	US-09-887-586A-57	Sequence 57, Appl
20	2031.5	62.5	2089	9	US-09-903-012-57	Sequence 57, Appl
21	2031.5	62.5	2089	10	US-09-900-797-57	Sequence 57, Appl
22	2031.5	62.5	2089	13	US-09-893-820-57	Sequence 57, Appl
23	2031.5	62.5	2089	14	US-10-041-007-23	Sequence 23, Appl
24	2031.5	62.5	2089	15	US-10-025-145A-5	Sequence 5, Appl
25	1295.5	39.8	1865	9	US-09-887-586A-47	Sequence 47, Appl
26	1295.5	39.8	1865	9	US-09-903-012-47	Sequence 47, Appl
27	1295.5	39.8	1865	10	US-09-900-797-47	Sequence 47, Appl
28	1295.5	39.8	1865	13	US-09-893-820-47	Sequence 47, Appl
29	1295.5	39.8	1865	14	US-10-041-007-17	Sequence 17, Appl
30	1250.5	38.2	1967	15	US-10-025-145A-17	Sequence 17, Appl
31	1240.5	38.2	2700	9	US-09-887-586A-43	Sequence 43, Appl
32	1240.5	38.2	2700	9	US-09-903-012-43	Sequence 43, Appl
33	1240.5	38.2	2700	10	US-09-900-797-43	Sequence 43, Appl
34	1240.5	38.2	2700	12	US-10-041-018-363	Sequence 363, App
35	1240.5	38.2	2700	13	US-09-893-820-43	Sequence 43, Appl
36	1231	37.9	1785	9	US-09-887-586A-49	Sequence 49, Appl
37	1231	37.9	1785	9	US-09-903-012-49	Sequence 49, Appl
38	1231	37.9	1785	10	US-09-900-797-49	Sequence 49, Appl
39	1231	37.9	1785	13	US-09-893-820-49	Sequence 49, Appl
40	1231	37.9	1785	14	US-10-041-007-19	Sequence 19, Appl
41	1187	36.5	2424	9	US-09-887-586A-45	Sequence 45, Appl
42	1187	36.5	2424	9	US-09-903-012-45	Sequence 45, Appl
43	1187	36.5	2424	10	US-09-900-797-45	Sequence 45, Appl
44	1187	36.5	2424	13	US-09-893-820-45	Sequence 45, Appl
45	1187	36.5	2424	14	US-10-041-007-14	Sequence 14, Appl

## ALIGNMENTS

RESULT 1  
US-10-025-145A-64  
; Sequence 64, Application US/10025145A  
; Publication No. US20030175861A1  
; GENERAL INFORMATION:  
; APPLICANT: Croteau, Rodney B.  
; APPLICANT: Bohlmann, Joerg  
; APPLICANT: Steele, Christopher L.  
; APPLICANT: Phillips, Michael A.  
; TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)  
; FILE REFERENCE: WSUR118414  
; CURRENT APPLICATION NUMBER: US/10/025,145A  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: US 09/360,545  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: PCT/US98/14528  
; PRIOR FILING DATE: 1998-07-10  
; PRIOR APPLICATION NUMBER: US 60/052,249  
; PRIOR FILING DATE: 1997-07-11  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 64  
; LENGTH: 2013

TYPE: DNA  
ORGANISM: Abies Grandis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (36) ..(1889)  
OTHER INFORMATION:  
US-10-025-145A-64

Alignment Scores:  
Pred. No.: 0 Length: 2013  
Score: 3251.00 Matches: 618  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 15 Gaps: 0

US-10-025-145A-65 (1-618) x US-10-025-145A-64 (1-2013)

QY 1 MetAlaLeuSerIleThrProLeuValSerArgSerCysLeuSerSerHisGlu 20  
DB 36 ATGGCTCTCTTCTATTAATCCGCTGGTTTCAGGTGTCCTCAGTCTTCTCATGAG 95  
QY 21 IleLysAlaLeuArgThrIleProThrLeuGlyIleCysArgProGlyLysSerVal 40  
DB 96 ATTAAGGCTCTCGTAGAACAAATCCCACTCTTGGAACTCTGACGGCGGGGAAATCCGTC 155  
QY 41 AlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAspSerValGlnArgArg 60  
DB 156 GCGCATTCCTAATAACATGTTGTTGACAAGCGTCGCACTACTACTGTTGACAGACGC 215  
QY 61 ValGlyAsnTyrHisSerAsnLeuTyrAspAspPheIleGlnSerLeuIleSerThr 80  
DB 216 GTGGGCACTATCAITTCACACCTGTGGACGATGATTTTCATACAGTCTCTGATCAACG 275  
QY 81 ProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIleGlyValLysAsp 100  
DB 276 CCTTATGGACACCTGATTACCGGGAACGCTGTCGACAGACTTATTTGGGGAAGTAAAGGAT 335  
QY 101 IleMetPheAsnPheLysSerLeuGluAspGlyGlyAsnAspLeuLeuGlnArgLeu 120  
DB 336 ATAATGTCAATTTCAAGTCGCTGGAGATGGAGGCAATGATCTCTTCAACGACTTTTG 395  
QY 121 LeuValAspAspValGluArgLeuGlyIleAspArgHisPheLysLysGlyIleLysThr 140  
DB 396 CTGGTCGATGACGTTGACGCTTTGGGAAATCGACAGGCAATTTCAAAAAGAGATAAAACG 455  
QY 141 AlaLeuAspTyrValAsnSerTyrTrpAsnGluLysGlyIleGlyCysGlyArgGluSer 160  
DB 456 GCACCTGATTATGTTAAACAGTTATTGGAACGAAAAGGCATTGGATCTGGGAGGGAGAT 515  
QY 161 ValValThrAspLeuAsnSerThrAlaLeuGlyLeuArgThrLeuArgLeuHisGlyTyr 180  
DB 516 GTTGTGATGACCTCAACTCAACCGCTTGGGGCTTGGAACTCTCCGACCTACCGGATAC 575  
QY 181 ThrValSerSerAspValLeuAsnValPheLysAspLysAsnGlyGlnPheSerSerThr 200  
DB 576 ACTGTGCTTCAGATGTTTGAACGTTTAAAGACAAAATGGCAATTTTCTCCCACT 635  
QY 201 AlaAsnIleGlnIleGluGlyIleArgGlyValLeuAsnLeuPheArgAlaSerLeu 220  
DB 636 GCCAATATTTCAGATAGAGGAGAGATTAGAGCGCTTCTCAATTTATTTCAGGGCTCCCTC 695  
QY 221 ValAlaPheProGlyGluLysValMetAspGluAlaGluThrPheSerThrLysTyrLeu 240  
DB 696 GTCCGCTTTTCCGGCGGAAGATTATGATGAGACCTGAACATTTCTTACAAAATATTA 755  
QY 241 ArgGluAlaLeuGlnLysIleProAlaSerSerIleLeuSerLeuGluIleArgAspVal 260  
DB 756 AGAAGAGCCCTGCAAAAGATTCCGGCATCCAGTATATCTTCTCACTAGAGATACGGGACGTT 815  
QY 261 LeuGluTyrGlyTyrHisThrAsnLeuProArgLeuGluAlaArgAsnTyrMetAspVal 280  
DB 816 CTGGATATGTTGGCACACCAATTTGCCAGCTTGGGAAGCAAGGAATTTACATGACGTC 875

QY 281 PheGlyGlnHisThrLysAsnLysAsnAlaAlaGluLysLeuLeuGluLeuAlaLysLeu 300  
DB 876 TTTGGACAGCACACTAAATAAAGAACGCGCGGAGAACTTTTAGAACTTCGAAATTCG 935  
QY 301 GluPheAsnIlePheHisSerLeuGlnGluArgGluLeuLysHisValSerArgTyrTrp 320  
DB 936 GAAATCAATATATTTCACTCTCTTCAAGAGAGAGAGTTAAACAATGTTTCCCGATCGTGG 995  
QY 321 LysAspSerGlySerProGluMetThrPheCysArgHisArgHisValGluTyrTrpAla 340  
DB 996 AAAGACTCGGGTCTCTCGAGATGACCTTCTGTGACATCGCTACAGTGGAAATACGCT 1055  
QY 341 LeuAlaSerCysIleAlaPheGluProGlnHisSerGlyPheArgLeuGlyPheThrLys 360  
DB 1056 TTGGCTTCTCGCATTCGTTGCGCTCAACATCTCGATTTCGACTCGGCTTTCACCAAG 1115  
QY 361 MetSerHisLeuIleThrValLeuAspAspMetTyrAspValPheGlyThrValAspGlu 380  
DB 1116 ATGCTCTCATCTTATCACGGTCTTGTACGACATGTACGAGCTCTTCGGCACAGTAGACGAG 1175  
QY 381 LeuGluLeuPheThrAlaThrIleLysArgTyrAspProSerAlaMetGluCysLeuPro 400  
DB 1176 CTGGAACTCTTTCACAGCGACATTAAGATGGATCCGTCGCGATGGAAATGCTTCCA 1235  
QY 401 GluTyrMetLysGlyValTyrMetValTyrHisThrValAsnGluMetAlaArgVal 420  
DB 1236 GAATATATGAAAGGAGTGTACATGATGTTTATCACCGCTAAATGAAATGGCTCGAGTG 1295  
QY 421 AlaGluLysAlaGlnGlyArgAspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCys 440  
DB 1296 GCAGAGAAGGCTCAAGCGCGAGACACGCTCAACTATGCAAGACAGGCTTGGGAGGCTGT 1355  
QY 441 PheAspSerTyrMetGlnGluAlaLysTyrIleAlaThrGlyTyrLeuProThrPheGlu 460  
DB 1356 TTTGATTCGTATATGAGGAAGCAAGTGGATCGCCACTGGTTATCTGCCCACTGGTTGAG 1415  
QY 461 GluTyrLeuGluAsnGlyLysValSerSerAlaHisArgProCysAlaLeuGlnProIle 480  
DB 1416 GAGTACTTGAGAAACGGGAAAGTTAGCTCTGCTCATCGCCCATCGCAGCTGCAACCCATT 1475  
QY 481 LeuThrIleAspIleProPheProAspHisIleLeuLysGluValAspPheProSerLys 500  
DB 1476 CTGACGTTGACATCCCTTCTGATCATCTCAAGGAAGTTGACTTCCCATCGAAG 1535  
QY 501 LeuAsnAspLeuIleCysIleIleLeuArgLeuArgGlyAspThrArgCysTyrLysAla 520  
DB 1536 CTCAATGACTTGATATGATCATCTTCGATTAAAGAGTGATACACGGTGCTCAACAGCA 1595  
QY 521 AspArgAlaArgGlyGluGluAlaSerSerIleSerCysTyrMetLysAspAsnProGly 540  
DB 1596 GACAGGCGCGTGGAGAAAGCTTCTGCTATATCATGTTATATGAAGACAAATCCCTGGA 1655  
QY 541 LeuThrGluGluAspAlaLeuAsnHisIleAsnPheMetIleArgAspAlaIleArgGlu 560  
DB 1656 TTAACGGAAGAGATGCTCTGAATCATATCATCTTATGATCAGGAGCGCAATCAGAGAA 1715  
QY 561 LeuAsnTrpGluLeuLysProAspAsnSerValProIleThrSerLysLysHisAla 580  
DB 1716 TTAATTTGGGAGCTTCTAAAGCCAGACACACAGTGTTCCTCATCTTCCAAAGAACACGCA 1775  
QY 581 PheAspIleSerArgValTyrHisGlyTyrArgTyrArgAspGlyTyrSerPheAla 600  
DB 1776 TTTGACATAAGCAGATTTCGCTACAGATACAGATACAGATGCTTACAGCTTTGCC 1835  
QY 601 AsnValGluThrLysSerLeuValMetArgThrValIleGluProValProLeu 618  
DB 1836 AACGTTGAAACAAACAGTTTGGTGATGAGAACCGCTCATTTGAACCTGTGCTTTG 1889

RESULT 2  
US-09-887-586A-19  
; Sequence 19, Application US/09887586A  
; Patent No. US20020094556A1

## GENERAL INFORMATION:

; APPLICANT: Chappell, Joseph  
 ; APPLICANT: No. US20020094556A11, Joseph P.  
 ; APPLICANT: Starks, Courtney M.  
 ; APPLICANT: Manna, Kathleen R.  
 ; TITLE OF INVENTION: SYNTHASES  
 ; FILE REFERENCE: 07678-025001  
 ; CURRENT APPLICATION NUMBER: US/09/887,586A  
 ; CURRENT FILING DATE: 2001-06-22  
 ; PRIOR APPLICATION NUMBER: 09/398,395  
 ; PRIOR FILING DATE: 1999-09-17  
 ; PRIOR APPLICATION NUMBER: 60/130,628  
 ; PRIOR FILING DATE: 1999-04-22  
 ; PRIOR APPLICATION NUMBER: 60/150,262  
 ; PRIOR FILING DATE: 1999-08-23  
 ; NUMBER OF SEQ ID NOS: 58  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 19  
 ; LENGTH: 2018  
 ; TYPE: DNA  
 ; ORGANISM: Abies grandis  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (6)...(1889)  
 ; OTHER INFORMATION: pinene synthase  
 US-09-887-586A-19

## Alignment Scores:

Pred. No.: 8,09e-279 Length: 2018  
 Score: 2453.00 Matches: 475  
 Percent Similarity: 84.63% Conservative: 59  
 Best Local Similarity: 75.28% Mismatches: 81  
 Query Match: 75.45% Indels: 16  
 DB: 9 Gaps: 6

US-10-025-145A-65 (1-618) x US-09-887-586A-19 (1-2018)

QY 1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15  
 DB 6 ATGGCTCTAGTTTCTACCGCACCGTGGCTTCCAAATCATGCTGCACAAATCGTTGATC 65  
 QY 16 SerSerHisGluIleLysAlaLeuArgThrIleProThrLeuGlyIleCysArg 35  
 DB 66 AGTTCTACCCATGAGCTTAAAGCTCTCTAGAACAAATTCAGCTCTAGGAATGATGATAGG 125  
 QY 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55  
 DB 126 CGAGGGAAATCTATCACTCTCTCCATCAGCATGAGCTCTACACCGTTTGAACCGATGAT 185  
 QY 56 SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTyrAspAspPheIleGln 75  
 DB 186 GGTGTAGAACCGCATGGGGCATTTCCATTCACACCTCTGGGACGATGATGTATACAG 245  
 QY 76 SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95  
 DB 246 TCTTTA---CCAACGGCTTATGAGGAAAAATCGTACCTGGAGCTGTGAGAAATCGATC 302  
 QY 96 GlyGluValLysAspIleMetPheAsnPhelLysSerLeuGluAspGlyGly----- 112  
 DB 303 GGGGAAGTAAGAAC---ATGTTCAATTCGATGTCTTAGAAGATGAGAGTAAATGAGT 359  
 QY 113 -----AsnAspLeuGlnArgLeuLeuValAspValGluArgLeuGlyIle 130  
 DB 360 CCGCTCAATGATCTCATTCACGCTTTGGATTGTTCGACACGCTTGAACGTTTGGGGATC 419  
 QY 131 AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150  
 DB 420 CATAGACATTTCAAAGATGAGATAAATCGCGCTTGTATATGTTTACAGTTATTTGGGGC 479  
 QY 151 GluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170  
 DB 480 GAAAAATGGCATCGGATCGGGGAGGAGATGTTGTTACTGATCTGAATCAACTCACTCGCTG 539

QY 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190  
 DB 540 GGGCTTCGAACCTACGACTACACGATACCGGTGTCTTCAGATGCTTTTCAAGAGCTTTC 599  
 QY 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGluLeuArg 210  
 DB 600 AAAGCCAAATGGCAGTTTTCTGCTCTGAAATATTCAGACAGATGAAGAGATCAGA 659  
 QY 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp 230  
 DB 660 GCGTTCCTGAATTTATTCGGGCTTCCTTCATTCCTTCAGGGGAGAAAAATTATGAT 719  
 QY 231 GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer 250  
 DB 720 GAGGCTGAAATCTCTTACCAATATTAAAGAGAGCCCTGCAAAAGATTCCGGTCTCC 779  
 QY 251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTyrHisThrAsnLeuPro 270  
 DB 780 AGT---CTTTCGCGAGATCGGGACGTTTTGGAAATATGTTGGCACACATATTTCGCG 836  
 QY 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsnLysAsnAla 290  
 DB 837 CGATTGGAAGCAGGAATTACATCCAGTCTTTGGACAGGACACTGAGAACGAGATCA 896  
 QY 291 -----AlaGluLysLeuLeuAlaLysLeuGluPheAsnIlePheHisSer 307  
 DB 897 TATGTGAAGAGCAAAAACTTTAGAACTCGCAAAATTTGGAGTTCAACATCTTTCAATCC 956  
 QY 308 LeuGlnGluArgGluLeuLysHisValSerArgTyrTrpLysAspSerGlySerProGlu 327  
 DB 957 TTACAAAAGAGGGAGTTAGAAAGTCTGTGATGTTGGTGAAGAAATCGGGTTTCTCTGAG 1016  
 QY 328 MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347  
 DB 1017 ATGACCTTCGCCGACATCGTCACTGGAATACTACATTTGGCTTCTCGATTCGGTTC 1076  
 QY 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal 367  
 DB 1077 GAGCTCAACATCTCGGATTCAGACTCGCTTGCAGAGCGTGTCTTATCATCGGTT 1136  
 QY 368 LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387  
 DB 1137 CTTCACGATATGTACGACACCTTCGGCACATAGACGAGCTGGAACCTCTTCACAGCGCA 1196  
 QY 388 IleLysArgTyrAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407  
 DB 1197 ATGAAGAGATGGATCGCTCTCGATAGATTGCCTTCAGATATATATGAAGAGGTGATC 1256  
 QY 408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427  
 DB 1257 ATAGCGGTTTACGACACCGTAAATGAAATGGCTCGAGAGGAGAGGCTCAAGGCCGA 1316  
 QY 428 AspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGlu 447  
 DB 1317 GATACGCTCATATGCTCGGAAAGCTTGGGAGGCTTATATTGATTCTGTATATGCAAGAA 1376  
 QY 448 AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys 467  
 DB 1377 GCAAGGTGGATCGCCACTGGTTACTTCCCTCTTCTGATGAGTACTACGAGAAATGGGAA 1436  
 QY 468 ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe 487  
 DB 1437 GTTAGCTGTGTCTATCGCATATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTC 1496  
 QY 488 ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle 507  
 DB 1497 CCTGATCATATCTCAAGGAGTTGATCTCCCATCAAAGCTTAAACGACTTGGCATGTGCC 1556  
 QY 508 IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu 527  
 DB 1557 ATCCTTCGATTACGAGGTGATACGGGTGCTACAGCGGACAGGGCTCTGTTGGAGAGAA 1616  
 QY 528 AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu 547

```

Db      1617 GCTTCTCTATATCATGTTATATGAAGCAATCTCGGAGTATCAGAGGAGATGCTCTC 1676
QY      548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLeuLys 567
Db      1677 GATCATATCAACGCCATGATCAGTGACGTAATCAAAAGGATTAAATTGGGAATCTCTCAA 1736
QY      568 ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTrp 587
Db      1737 CCAGACATCAATGTTCCCATCTCGGCAAGAAACATGCTTTTGACATCGCCAGAGCTTTC 1796
QY      588 HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu 607
Db      1797 CATTAGCGCTACAAATACCGAGACGGCTACAGGTTGCCAACGTTGAAACGAGAGATTG 1856
QY      608 ValMetArgThrValIleGluProValProLeu 618
Db      1857 GTACAGAGAACCTCTCTGTAATCTGTGCTTTG 1889

RESULT 3
US-09-903-012-19
; Sequence 19, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6)...(1889)
; OTHER INFORMATION: pinene synthase
US-09-903-012-19

Alignment Scores:
Pred. No.:      8,09e-279      Length:      2018
Score:          2453.00      Matches:      475
Percent Similarity: 84.63%      Conservative: 59
Best Local Similarity: 75.28%      Mismatches: 81
Query Match:      75.45%      Indels:      16
DB:              9          Gaps:         6

US-10-025-145a-65 (1-618) x US-09-903-012-19 (1-2018)
QY      1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15
Db      6 ATGGCTCTAGTTTCTACCGCACCGTGGCTTCCAAATCATGCTTGCACAAATCGTTGATC 65
QY      16 SerSerSerHisGluIleLysAlaLeuArgArgThrIleProThrLeuGlyIleCysArg 35
Db      66 AGTCTACCCATGAGCTTAAGCTCTCTAGAACAAATTCAGCTCTAGGAATGAGTAGG 125
QY      36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55
Db      126 CGAGGGAATCTACTACTCTCTCCATCAGCATGAGCTTACCACCGTTGTAAACCGATGAT 185

```

```

QY      56 SerValGlnArgValGlyAsnTyrHisSerAsnLeuTirAspAspPheIleGln 75
Db      186 GGTGTACGAGACGATGGCGATTTCCATTCCACCTCTGGGACGATGATGTACACAG 245
QY      76 SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95
Db      246 TCTTTA---CCAGCGCTTATCAGGAAAAATCGTACCTGGAGCGTCTGAGAACTGATC 302
QY      96 GlyGluValLysAspIleMetPheAsnPheLysSerLeuGluAspGlyGly----- 112
Db      303 GGGGAAGTAAAGAAC---ATGTTCAATTCGATGTCTATTAGAGATCGAGGTTAATGAGT 359
QY      113 -----AsnAspLeuLeuGlnArgLeuLeuValAspAspValGluArgLeuGlyIle 130
Db      360 CGCTCAATGATCTCATTCAGCGCTTTGGATGTTCGACAGCTTCAAGCTTTGGGATC 419
QY      131 AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150
Db      420 CATAGACATTTCAAGATGAGATAAAATCGGCGCTTGATTATGTTTACAGTTATTGGGGC 479
QY      151 GluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170
Db      480 GAAATGCGCATCGGATCGGCGAGGAGTGTGTTACTGATCTGAACCTCAACTGCGTTG 539
QY      171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190
Db      540 GGGCTTCGAACCTTCAGACTACAGGATACCCGGTGTCTTCAGATGTTTCAAAGCTTTC 599
QY      191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnLeuGluGlyIleArg 210
Db      600 AAAGGCCAAAAATGGCAGTTTTCTCTGCTCTGAAAAATATTCAGACAGATGAAGAGATCAGA 659
QY      211 GlyValLeuAsnLeuPheArgAlaSerLeuAlaAlaPheProGlyGluLysValMetAsp 230
Db      660 GCGCTTCTGAATTTATTCGGGCTCTCCCTCATTTGCTTTCCAGGGGAGAAAAATATGAT 719
QY      231 GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer 250
Db      720 GAGCTGAAATCTCTCTACCAATATTTAAAGAGAGCCCTGCAGAAAGATTCGGGTCTCC 779
QY      251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTyrHisThrAsnLeuPro 270
Db      780 AGT---CTTTTCGGCAGAGATCGGGGACGTTTTTGAATATGTTGGCACACATATTTCCCG 836
QY      271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsnLysAsnAla 290
Db      837 CGATTGGAACGAAGGAATTACATCCAAAGTCTTTGGACAGGACACTGAGACACGAACTCA 896
QY      291 -----AlaGluLysLeuLeuGluAlaLysLeuGluPheAsnIlePheHisSer 307
Db      897 TATGTGAAGAGCAAAAAACTTTTAGAACTCGCAAAATTTGAGATTCAACATCTTTCAATCC 956
QY      308 LeuGlnGluArgGluLeuLysHisValSerArgTrpTrpLysAspSerGlySerProGlu 327
Db      957 TTCAAAAGAGGGAGTTAGAAAGTCTGCTCAGATGGTGGAAAGATCGGGTTTTCCTGAG 1016
QY      328 MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347
Db      1017 ATGACCTTCTGCCGACATCGTCACTGGAATACTACATTTGGCTTCTCGCATTCGCTTC 1076
QY      348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal 367
Db      1077 GAGCTCAACATCTCGATTTCAGACTCGGCTTTGCCAAGAGCTGTCTATCTTATCACGGTT 1136
QY      368 LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387
Db      1137 CTTGACCATATGTACGACACCTTCGGCACAGTAGACGCTGGAACTCTTTCACAGCCACA 1196
QY      388 IleLysArgTrpAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407
Db      1197 ATGAAGAGATGGGATCCGCTCTCGATAGATTGCTTCCAGAAATATATGAAAGAGATGTAC 1256

```

```

QY 408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427
Db 1257 ATAGCGTTTACGACACCGTAATGAATGGCTCGAGAGCGAGGAGGCTCAAGGCCGA 1316
QY 428 AspThrLeuAsnTyrAlaArgGlnAlaTyrGluAlaCysPheAspSerTyrMetGlnGlu 447
Db 1317 GATACGCTCACATATGCTCGGAAGCTTGGAGGCTTATATTGATTCGTATATGCAAGAA 1376
QY 448 AlaLysTyrPheAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys 467
Db 1377 GCAAGGTGATCGCCACTGTTACTCTGCTGCTCTTGTGAGTACTTACGAGAATGGAAA 1436
QY 468 ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe 487
Db 1437 GTTAGCTGGTTCATCGCATATCCGCAATTCACCAATTCGACAAATGACATCCCTTT 1496
QY 488 ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle 507
Db 1497 CCTGATCATATCTCAAGGAAGTTGACTTCCCATCAAAGCTTAAACGACTTGGCATGTGCC 1556
QY 508 IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspAlaArgGlyGluGlu 527
Db 1557 ATCCCTTCGATTACGAGGTGATACGGGTGCTACAAAGCGGACAGGGCTCGTCGAGAGAA 1616
QY 528 AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu 547
Db 1617 GCTTCCTATATCATGTTATATGAAGACAATCCTCGAGTATCAGAGGAATGCTCTC 1676
QY 548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTyrGluLeuLys 567
Db 1677 GATCATATCAACGCCATGATCAGTCAGCTAATCAAGGATTAATTTGGAACTTCTCAA 1736
QY 568 ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTyr 587
Db 1737 CCAGACATCAATGTTCCTATCTCGCGAAGAAACATGCTTTTGACATCGCCAGAGCTTTC 1796
QY 588 HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu 607
Db 1797 CATTAGCGTACAATACCGAGACGGCTACAGCGTTGCAACGTTGAAACGAGAGTTTG 1856
QY 608 ValMetArgThrValIleGluProValProLeu 618
Db 1857 GTCACGAGAACCTCTCTTGATCTGTGCTTTG 1889

```

RESULT 4

```

US-09-900-797-19
; Sequence 19, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6)...(1889)
; OTHER INFORMATION: pinene synthase

```

```

US-09-900-797-19
Alignment Scores:
Pred. No.: 8,09e-279 Length: 2018
Score: 2453.00 Matches: 475
Percent Similarity: 84.63% Conservative: 59
Best Local Similarity: 75.28% Mismatches: 81
Query Match: 75.45% Indels: 16
DB: 10 Gaps: 6
US-10-025-145a-65 (1-618) x US-09-900-797-19 (1-2018)

```

```

QY 1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15
Db 6 ATGGCTCTAGTTTCTACCGCACCGTGGCTTCCAAATCATGCTGCACAAATCGTGTATC 65
QY 16 SerSerSerHisGluIleLysAlaLeuArgThrIleProThrLeuGlyIleCysArg 35
Db 66 AGTTCTACCCATGAGCTTAAGGCTCTCTCTAGAACAAATCCAGCTCTAGGAATGAGTAGG 125
QY 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55
Db 126 CGAGGGAATTCATCATCTCTTCCATCAGCATGAGCTCTACCCAGCTTTGTAACCATGAT 185
QY 56 SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTyrAspAspPheIleGln 75
Db 186 GGTGTACGAGACGCGATGGCGGATTTCCATTCACACCTCTGGGACGATGATGTATACAG 245
QY 76 SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95
Db 246 TCTTTA---CCAACGGCTTATGAGGAAAAATCGTACCTGGAGCGTGTGAGAAAACTGATC 302
QY 96 GlyGluValLysAspIleMetPheAsnPhelysSerLeuGluAspGlyCly----- 112
Db 303 GGGGAGTAAAGAAC---ATGTTCAATTCGATGTCATTAGAGATGAGAGTAAATGAGT 359
QY 113 -----AsnAspLeuLeuGlnArgLeuLeuValAspAspValGluArgLeuGlyIle 130
Db 360 CCGCTCAATGATCTCATTCACCGCTTTGGATTGTGACAGCCTTGAACGTTGGGGATC 419
QY 131 AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150
Db 420 CATAGACATTTCAAAGATGAGATAAAATCGCGCTTGATTTATGTTTACAGTTATTGGGCG 479
QY 151 GluLysGlyIleGlyCysGlyArgGluSerValThrAspLeuAsnSerThrAlaLeu 170
Db 480 GAATAATGGCATCGGATCGGGAGGAGAGTGTGTTACTGATCTGAACCTCAACTCGCGTTG 539
QY 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190
Db 540 GGGCTTCGAACCTTACGACTACAGGATACCCGCTGTCTTCAGATGTTTCAAAGCTTTC 599
QY 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGluIleArg 210
Db 600 AAAGGCCAAAATGGCGAGTTTTCCTGCTCTCAAAATATTCAGACAGATGAAGATCAGA 659
QY 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp 230
Db 660 GCGCTTCTGAATTTATTCGCGGCTCCCTCATTCGCTTCCAGGGGAGAAAATATATGAT 719
QY 231 GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer 250
Db 720 GAGGCTGAATCTCTCTACCAATATTTAAAGAGCCCTGCAAAAGATTCGGGTCTCC 779
QY 251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTrpHisThrAsnLeuPro 270
Db 780 AGT---CTTTCCGAGAGATCGGGAGCGTTTGGAAATATGTTGGCACACATATTTCGCG 836
QY 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsnAla 290
Db 837 CGATTGGAGCAAGGAATTACATCCAAAGTCTTTGGACAGGACACTGAGAACAGAGTCA 896
QY 291 -----AlaGluLysLeuLeuGluLeuAlaLysLeuGluPheAsnIlePheHisSer 307

```

```

; APPLICANT: No. US20040053386A11, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6)...(1889)
; OTHER INFORMATION: pinene synthase
US-09-893-820-19

Alignment Scores:
Pred. No.: 8,09e-279 Length: 2018
Score: 2453.00 Matches: 475
Percent Similarity: 84.63% Conservative: 59
Best Local Similarity: 75.28% Mismatches: 81
Query Match: 75.45% Indels: 16
DB: 13 Gaps: 6

US-10-025-145A-65 (1-618) x US-09-893-820-19 (1-2018)
Qy 1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15
Db 6 ATGGCTCTAGTTCTTACCGCAGCGTTGGCTTCCAAATCATGCTGCACAAATCGTTGATC 65
Qy 16 SerSerSerHisGluLeuLeuAlaLeuArgArgThrIleProThrLeuGlyIleCysArg 35
Db 66 AGTTCTACCATCAGCTTAAGGCTCTCTAGAACAAATCCAGCTCTAGAAATGAGTAGG 125
Qy 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55
Db 126 CGAGGGAATCTATCACTCTCTCCATCAGCATGAGCTCTACCACGTTGTAACCGATGAT 185
Qy 56 SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIleGln 75
Db 186 GGTGTACGAAGACGATGGGCGATTTCCATTCCAACCTCTGGGACGATGATGTCATACAG 245
Qy 76 SerLeuLeuSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95
Db 246 TCTTTA---CCAAACGGCTTATGAGGAAAAATCGTACTCGGAGCGTGTGAGAAACTGATC 302
Qy 96 GlyGluValLysAspIleMetPheAsnPhelysSerLeuGluAspGlyGly----- 112
Db 303 GGGGAAGTAAGAAC---ATGTTCAATTCGATGTCTTAGAAGATGAGAGAGTAAATGAGT 359
Qy 113 -----AsnAspLeuLeuGlnArgLeuLeuValAspAspValGluArgLeuGlyIle 130
Db 360 CCGCTCAATGATCTCAATCAACGCGCTTTGGATTCTCGACACGCTTGAAACGTTTGGGGATC 419
Qy 131 AspArgHisPheLysLysGluLeuLeuThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150
Db 420 CATAGACATTTCAAAGATGAGATAAATCGCGCTTGNATGATGTTTACAGTTATTGGGCGC 479
Qy 151 GluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170
Db 480 GAAATGCGCATCGGATCGGAGGAGAGAGTGTGTACTGATCTGTAACCTCAACTCACTGCGTTG 539

RESULT 5
US-09-893-820-19
; Sequence 19, Application US/09893820
; Publication No. US20040053386A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
```

```
QY 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190
DB 540 GGGCTTCGACCCCTACGACTACCGATACCGGGTGTCTTCAGATGTTTCAAAGCGTTTC 599
QY 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGlyGluIleArg 210
DB 600 AAAGGCCAAATGGGCAGTTTCTCTGCTGCTGAAATATTCACAGACATGAAGATCAGA 659
QY 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp 230
DB 660 GCGGTTCTGAATTTATTCGGGCCCTCCCTCATTTGCTTTCCAGGGGAGAAAATTTAGGAT 719
QY 231 GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer 250
DB 720 GAGGCTGAAATCTTCTACCAAAATATTTAAAGGAAGCCCTCGAAAGATTCGGGTCTCC 779
QY 251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTrpHisThrAsnLeuPro 270
DB 780 AGT---CTTTCGGAGAGATCGGGCAGCTTTTGGAAATATGTTGGCACACATATTGGCG 836
QY 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsnLysAsnAla 290
DB 837 CGATTGGAAGCAAGAAATTAATCCCAAGTCTTTGGACAGGACACTCAGAAACACGAAGTCA 896
QY 291 -----AlaGluLysLeuLeuGluLeuAlaLysLeuGluPheAsnIlePheHisSer 307
DB 897 TATGTGAAGAGCAAAACTTTTAGAATCGCAAAATTTGGAGTTCAACATCTTTCAAATCC 956
QY 308 LeuGlnGluArgGluLeuLysHisValSerArgTrpTrpLysAspSerGlySerProGlu 327
DB 957 TTACAAAGAGGGAGTTAGAAAGTCTGTCAGATGTCGGAAGATCGGGTTTCTCGAG 1016
QY 328 MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347
DB 1017 ATGACCTTCTCCGACATCGTCAGCTGGAATACTACATTTGGCTTCTCGCATTCGCGTTC 1076
QY 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal 367
DB 1077 GAGCTCAACATTTCTGNTTCAGACTCGGCTTTCGCAAGAGCTGTCTATCATCAGCGTT 1136
QY 368 LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387
DB 1137 CTTGACGATATGTACGACACCTTCGGCACAGTAGACGAGTGGAACTCTTCACAGCGACA 1196
QY 388 IleLysArgTrpAspProSerAlaMetClyLeuProGluTyrMetLysGlyValTyr 407
DB 1197 ATGAAGAGATGGGATCCGCTCCGATGATTCGCTTCAGAAATATATCAAAAGGAGTGTAC 1256
QY 408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427
DB 1257 ATAGCGGTTTACGACACCGTAAATGAAATGGCTCGAGGGCAGAGGAGCTCAAGCCGA 1316
QY 428 AspThrLeuAsnTyrAlaArgGlnAlaTrpGluAlaCysPheAspSerTyrMetGlnGlu 447
DB 1317 GATACGCTCATATGCTCGGGAAGCTTGGAGGCTTATATGATTGCTATATCAAGAA 1376
QY 448 AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys 467
DB 1377 GCRAAGTGGATCGCCATCGGTACTCGCTCCCTTCGATGAGTACTACGAGAATGGGAAA 1436
QY 468 ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe 487
DB 1437 GTTAGCTGTGTCATCGCATATCGCATTCGAACCCATTTCGACATGGACATCCCCCTTT 1496
QY 488 ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle 507
DB 1497 CCTGATCATATCCTCAAGGAAGTTGACTTCCCATCAAAGCTTAACGACTTGGCATGTGCC 1556
QY 508 IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu 527
DB 1557 ATCCTTCGATTACGAGTGATACGCGTGTCTACAAGCGGACACAGGGCTCGTGGAGAAGAA 1616
QY 528 AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluAspAlaLeu 547
```

```
DB 1617 GCTTCCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTC 1676
QY 548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLeuLys 567
DB 1677 GATCATATCAACGCCCATGATCATGACGTGTAATCAAAAGGATTAAATTTGGGAACCTTCTCAA 1736
QY 568 ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTrp 587
DB 1737 CCAGACATCAATGTTCCCATCTCTCGGCGAAGAAACATGCTTTTGACATGCCAGAGCTTTC 1796
QY 588 HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu 607
DB 1797 CATTCAGGCTACAAATACCGACGCTACAGCGTTGCCACGTTGCAACAGGAGGTTTG 1856
QY 608 ValMetArgThrValIleGluProValProLeu 618
DB 1857 GTCACGAGAACCTCTCTGAATCTGTGCTTTG 1889

RESULT 6
US-10-041-007-21
; Sequence 21, Application US/10041007
; Publication No. US20020164736A1
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seichi P.T.
; APPLICANT: Schepmann, Hala G
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase
; FILE REFERENCE: P02081US1
; CURRENT APPLICATION NUMBER: US/10/041,007
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/259,881
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Abies grandis
US-10-041-007-21

Alignment Scores:
Pred. No.: 8,09e-279 Length: 2018
Score: 2453.00 Matches: 475
Percent Similarity: 84.63% Conservative: 59
Best Local Similarity: 75.28% Mismatches: 81
Query Match: 75.45% Indels: 16
DB: 14 Gaps: 6

US-10-025-145A-65 (1-618) x US-10-041-007-21 (1-2018)
QY 1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15
DB 6 ATGGCTCTAGTTTCTACCGCACCGTTCGGCTTCCAAATCATGCTCGCACAAATCGTTGATC 65
QY 16 SerSerSerHisGluIleLysAlaLeuArgArgThrIleProThrLeuGlyIleCysArg 35
DB 66 AGTTTACCCTACGCTTAAGGCTCTCTCTAGAACAAATTCAGCTCTAGGAAAGAGTAGG 125
QY 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55
DB 126 CGAGGGGAAATCTATCTCTCTCCATCAGCATGAGCTCTACCACCGTTGTAACCGATGAT 185
QY 56 SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIleGln 75
DB 186 GGTGTACGAGAGCGCATGGCGGATTTCCATTCCAACTCTGGGACCATGATGTCATACAG 245
QY 76 SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95
DB 246 TCTTTA---CCAACGGCTTATGAGGAAAAATCGTACTCGGAGCGTCTGAGAAACTGATC 302
QY 96 GlyGluValLysAspIleMetPheAsnPheLysSerLeuGluAspGlyGly----- 112
DB 303 GGGGAAAGTAAAGAAC---ATGTTTCAATTCGATGCTTAGAAGATGAGAGGTTAATGAGT 359
```

QY	113	-----AsnAspLeuLeuGlnArgLeuLeuValAspAspValGluArgLeuGlyIle	130	QY	468	ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe	487
Db	360	CGGCTCAATGATCTCATCAACGCCCTTTGGATTGTGCGAGCCCTTGAACGTTTGGGGATC	419	Db	1437	GTTAGCTGTGGTCAATCGCATATCCGCAATTCGCAACCACTTCTGACAAATGACATCCCTTT	1496
QY	131	AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn	150	QY	488	ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspIleuileCysIle	507
Db	420	CATAGACATTTCAAAGATGAGATAAATCGGCCCTTGATTATGTTTACAGATTATTGGGGC	479	Db	1497	CCTGATCATATCTCTCAAGGAAGTTGACTTCCCATCAAGCTTAAAGACTTGGCATGTGCC	1556
QY	151	GluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu	170	QY	508	IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu	527
Db	480	GAAATGGCATCGGATCGGGAGGAGAGTGTGTTACTGATCTGAACCTCAACTCGGTTG	539	Db	1557	ATCTCTTCATATACGAGGTGATACGGGTGCTCAAGAGCGGACAGGGCTGTGGAGAGAA	1616
QY	171	GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe	190	QY	528	AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu	547
Db	540	GGGCTTCGAACCTACGACTACACGGATACCGCGTGTCTTCAGATTTTTTCAAGGCTTTC	599	Db	1617	GCTTCTCTATATCATGTTATATGAAGACAATCTCTGGAGTATCAGAGGAAGATGCTCTC	1676
QY	191	LysAspLysAsnGlyGlnPheSerThrAlaAsnIleGlnIleGluGlyGluIleArg	210	QY	548	AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLys	567
Db	600	AAAGGCCAAATGGGCAGTTCCTCTCTGAAATATTTTCAGACAGATGAAGAGATCAGA	659	Db	1677	GATCATATCAACGCCATGATGATGACGTAATCAAGGATTAATTTGGGAACCTTCTCAA	1736
QY	211	GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp	230	QY	568	ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTrp	587
Db	660	GGGTTCTGAATTTATTCGGGCCCTCCCTCATTCCTTTCCAGGGGAGAAATATGGAT	719	Db	1737	CCAGACATCATATGTTCCCATCTCGGGAGAAACAATGCTTTTGACATCCCGAGGCTTC	1796
QY	231	GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer	250	QY	588	HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu	607
Db	720	GAGGCTGAATCTCTCTACCAATATTTAAAGAGCCCTGCAAAAGATTCGGGTCTCC	779	Db	1797	CATTACGGCTACAAATACCGAGCGGTACAGGTTGCCAACGTTGAAACGAGAGATTG	1856
QY	251	SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTrpHisThrAsnLeuPro	270	QY	608	ValMetArgThrValIleGluProValProLeu	618
Db	780	AGT---CTTTCCGAGAGATCGGGGACGTTTGGAAATATGTTGGACACATATTTGCCG	836	Db	1857	GTACAGAGAACCTCTTGAATCTGTGCCTTTG	1889
QY	271	ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLysAsnLysAsnAla	290	RESULT 7			
Db	837	CGATTGGAAGCAAGGAATATACATCCAGTCTTTGGACAGGACACTCAGAACACGAAGTCA	896	US-10-025-145A-3			
QY	291	-----AlaGluLysLeuLeuGluLeuAlaLysLeuGluPheAsnIlePheHisSer	307	; Sequence 3, Application US/10025145A			
Db	897	TATGTGAAGACAGCAAAACTTTTAGAACTCGCAAAATTTGGAGTTCAACATCTTCAATCC	956	; Publication No. US20030175861A1			
QY	308	LeuGluGluArgGluLeuLysHisValSerArgTyrTriLysAspSerGlySerProGlu	327	; GENERAL INFORMATION:			
Db	957	TTCAAAAGAGGAGTTAGAAAGTCTGGTCAGATGGTGGAAAGATCGGGTTTCTCTGAG	1016	; APPLICANT: Croteau, Rodney B.			
QY	328	MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe	347	; APPLICANT: Bohlmann, Joerg			
Db	1017	ATGACCTTCTGCCGACATCGTCACTGGAATACTACACTTTGGCTTCTGCAATGGCTTC	1076	; APPLICANT: Steele, Christopher L.			
QY	348	GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal	367	; APPLICANT: Phillips, Michael A.			
Db	1077	GAGCCTCAACATTTGGATTTCAGACTCGGCTTTGCCAAGCGTGTCACTTATTCACCGTT	1136	; TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)			
QY	368	LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr	387	; FILE REFERENCE: WSUR118414			
Db	1137	CTTGACGATATGACGACACCTTCGGCACAGTAGACGAGCTGGAACCTTTCACAGCGACA	1196	; CURRENT APPLICATION NUMBER: US/10/025,145A			
QY	388	IleLysArgTrpAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr	407	; CURRENT FILING DATE: 2002-06-28			
Db	1197	ATGAAGAGATGGGATCGCTCTCATAGATTGCTTCCAGAAATATATGAAGGAGTGTAC	1256	; PRIOR APPLICATION NUMBER: US 09/360,545			
QY	408	MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg	427	; PRIOR FILING DATE: 1999-07-26			
Db	1257	ATAGCGGTTTACGACACCGTAATATGAATGGCTCGAGAGGAGGAGGCTCAAGGCCGA	1316	; PRIOR APPLICATION NUMBER: PCT/US98/14528			
QY	428	AspThrLeuAsnTyrAlaArgGlnAlaArgPheGluAlaCysPheAspSerTyrMetGlnGlu	447	; PRIOR FILING DATE: 1998-07-10			
Db	1317	GATACGCTCATATGCTCGGAAGCTTGGAGGCTTATATTGATTTCGTATATGCAAGAA	1376	; PRIOR APPLICATION NUMBER: US 60/052,249			
QY	448	AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys	467	; PRIOR FILING DATE: 1997-07-11			
Db	1377	GCAAGTGGATCGGCACTGGTTTACCTGCCCTCTTGTATGAGTACTACGAGATGGGAAA	1436	; NUMBER OF SEQ ID NOS: 107			
				; SOFTWARE: PatentIn version 3.1			
				; SEQ ID NO 3			
				; LENGTH: 2018			
				; TYPE: DNA			
				; ORGANISM: Abies Grandis			
				; FEATURE:			
				; NAME/KEY: CDS			
				; LOCATION: (6)..(1892)			
				; OTHER INFORMATION:			
				US-10-025-145A-3			
				Alignment Scores:			
				Pred. No.: 8,09e-279	Length: 2018		
				Score: 2453.00	Matches: 475		
				Percent Similarity: 84.63%	Conservative: 59		
				Best Local Similarity: 75.28%	Mismatches: 81		
				Query Match: 75.45%	Indels: 16		
				DB: 15	Gaps: 6		
				US-10-025-145A-65 (1-618) x US-10-025-145A-3 (1-2018)			



; PRIOR FILING DATE: 1997-07-11  
 ; NUMBER OF SEQ ID NOS: 107  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 66  
 ; LENGTH: 2186  
 ; TYPE: DNA  
 ; ORGANISM: Abies Grandis  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (34)..(1923)  
 ; OTHER INFORMATION:  
 ;  
 US-10-025-145a-66

Alignment Scores:  
 Pred. No.: 1.97e-256 Length: 2186  
 Score: 2264.00 Matches: 435  
 Percent Similarity: 80.53% Conservative: 78  
 Best Local Similarity: 68.29% Mismatches: 98  
 Query Match: 69.64% Indels: 26  
 DB: Gaps: 9

US-10-025-145a-65 (1-618) x US-10-025-145a-66 (1-2186)

QY 1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu-----15  
 DB 34 ATGGCTCTGGTTTTCGCCGACCC-----AAATCTGCTGCCGACAAATCGTTGATC 84  
 QY 16 ---SerSerSerHisGluIleLeuAlaLeuArgThrIleProThrLeuGlyIleCys 34  
 DB 85 AGGTCTACTCATCATGAGCTCAAGCCTCTGCGAGAACCATCCCACTCTTGGAATGTGT 144  
 QY 35 ArgProGlyIleSerSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThr 54  
 DB 145 AGGGAGAGGAAATTTTCAACCTTCTGTGAGCAAGAGTTTGACCAACCGCTGTATCTGAT 204  
 QY 55 AspSerValGlnArgArgValGlyAsnTyrlHisSerAsnLeuTrpAspAspPheIle 74  
 DB 205 GATGGTCTACAAAGACGATGAGTGTGACTATCATTCACATCTCTGGACGACGATTTCTA 264  
 QY 75 GlnSerLeuIleSerThrProTyrlValAlaProAspTyrlArgGluValAlaAspArgLeu 94  
 DB 265 CAGTCTCTA---TCAGCGCTTATGCGGAGCGCTTCTTACCGAGAACGTCGAGAACTGT 321  
 QY 95 IleGlyGluValIleAspIleMetPheAsnPhelySerIleGluAspGlyGly-----112  
 DB 322 ATTGGGAGAGGAGAGG---ATGTTCAATTCATGCCATCGAGAGATGGGAATCAATG 378  
 QY 113 -----AsnAspLeuLeuGlnArgLeuLeuValAspAspValGluArgLeuGly 129  
 DB 379 AGTCCCTCCATGATCTTATTGAACGACTTTGGATGTCGATAGCGTTGAACGTTTGGGG 438  
 QY 130 IleAspArgHisPheIleValGlyGluIleLeuSerThrAlaLeuAspTyrlValAsnSerTyrlTrp 149  
 DB 439 ATTGATGACCATTTTCAAAAAGAGATAAATACAGCCCTTGATTAATGTTTACATTAATGG 498  
 QY 150 AsnGluIleValIleGlyCysGlyArgGlnSerValAlaThrAspLeuAsnSerThrAla 169  
 DB 499 AACGAAAGAGTATTGATGAGTGCAGTGAATGATGTTTTCCTGATGTCACATCCGACTGCC 558  
 QY 170 LeuGlyLeuArgThrLeuArgLeuHisGlyTyrlThrValIleSerSerAspValIleuAsnVal 189  
 DB 559 TCGGGGTTTCGAACCTTCGCTCACGAGATACAGTGTCTCTTCAGAGGTTTGAAGTA 618  
 QY 190 PheIleAspIleValIleGlyIlePheSerSerThrAlaAsnIleGlnIleGluIleGluIle 209  
 DB 619 TTTCAAGACCAAAAGTGGCAGTTTGCATTCCTCTCTAGTACAAA---GAGGAGACACATC 675  
 QY 210 ArgGlyValIleuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluIleValMet 229  
 DB 676 AGAACCGTTTCAATTTATATGCGGCTTCTTTCATTTGCTTCTCGGGAGAAATTTATG 735  
 QY 230 AspGluAlaGluThrPheSerThrIleTyrlLeuArgGluAlaLeuGlnIleIleProAla 249

DB 736 GAAGAGCTGAATTTTCTCTCAAGATATTGAAAGAACCGCTGCAAAAGATTCGCGTC 795  
 QY 250 SerSerIleLeuSerLeuGluIleArgAspValLeuGluIleTyrlGlyTrpHisSerAsnLeu 269  
 DB 796 TCCAGT---CTTTCACAAAGAAATGACTACACTTTGGAAATATGTTGGCACAAATATG 852  
 QY 270 ProArgLeuGluIleAlaArgAsnTyrlMetAspValPheGlyGlnHisThr-----285  
 DB 853 CCAAGATTGGAACCAAGAAATTACTTAGATGTTATTGGACATCCCTACAGTCATGCGCTC 912  
 QY 286 LysAsnLysAsnAla-----AlaGluIleLeuLeuGluIleuAlaIleLeuGlu 301  
 DB 913 AAGAAAGAAAGAGACGCAATATCTGACAGGAAAGCTTTTGAAGTTCGCAAAATTTGGAG 972  
 QY 302 PheAsnIlePheHisSerLeuGlnGluArgGluLeuIleHisValIleSerArgTrpTrpLys 321  
 DB 973 TTCAACATCTTTTCACTCTCTTCAACAGAGAGTTACAGTATCTCTCCAGATGGTGATA 1092  
 QY 322 AspSerGlySerProGluMetThrPheCysArgHisArgHisValGluTyrlValAlaLeu 341  
 DB 1093 CATTGGGTTTGCCTGAATGACCTTTGTGCGGATCGTCACAGTGAATCTACACCCCTG 1092  
 QY 342 AlaserCysIleAlaPheGluProGlnHisSerGlyPheArgLeuGlyPheThrIleMet 361  
 DB 1093 AGCTCTTGCAATGGAGCTGAGCCCAACATTCGATTCAGATTTGGGCTTTGCCAAAACG 1152  
 QY 362 SerHisLeuIleThrValLeuAspAspMetTyrlAspValPheGlyThrValAspGluLeu 381  
 DB 1153 TGTCATCTTTACACGGTTCGAGATATCTACACACCTTTCCGAGAGATGGATGAAATC 1212  
 QY 382 GluLeuPheThrAlaThrIleLysArgTrpAspProSerAlaMetGluCysLeuProGlu 401  
 DB 1213 GAACTTTCAACGAGGAGTAGAGATGGAATCGCTGCGAGAAAGAACCGCTCCCGAA 1272  
 QY 402 TyrlMetLysGlyValTyrlMetMetValTyrlHisThrValAsnGluMetAlaArgValAla 421  
 DB 1273 TATATGAAAGAAATCATATGATGCACTTACGAGCCCTTAAGTGAATGCGCGAGAGCA 1332  
 QY 422 GluLysAlaGlnGlyArgAspThrLeuAsnTyrlAlaArgGlnAlaTrpGluAlaCysPhe 441  
 DB 1333 GAGAAAGACAGGCGCAGACACCGCTCATATATGCTAGAAAGCGTTGGAAATTTATCTT 1392  
 QY 442 AspSerTyrlMetGlnGluAlaIleTyrlAlaThrGlyTyrlLeuProThrPheGluIle 461  
 DB 1393 GATTCGATACACAAGAGCAAGTGAATCGCCAGCGGTTATCTGCCAATCTTGAGAGAG 1452  
 QY 462 TyrlLeuGluAsnGlyIleValSerSerAlaHisArgProCysAlaLeuGlnProIleLeu 481  
 DB 1453 TACTTAGAAGCGCGAAGTTAGCTGTGTCATGTGCGACGCGCATTTGACACCCCTCTGT 1512  
 QY 482 ThrLeuAspIleProPheProAspHisIleLeuIleGluArgLeuArgGlyAspThrArgCysTyrlValAsp 501  
 DB 1513 ACATTTGACGATACCGCTTCTCGATGAGCTTGAAGGAAATAGATTTTCCATCGAGATTT 1572  
 QY 502 AsnAspLeuIleCysIleIleLeuArgLeuArgGlyAspThrArgCysTyrlValAsp 521  
 DB 1573 AATGATTTGGCATCTTCTCTTCTTACAGTGAAGAGTGCACACGATGTCACAAAGCAAGC 1632  
 QY 522 ArgAlaArgGlyGluGluAlaSerSerIleSerCysTyrlMetLysAspAspProGlyLeu 541  
 DB 1633 AGGACCGAGAGAGAGAGCGCTCAAGCATATCGTTCATCATGAAGACATCCCGGATTA 1692  
 QY 542 ThrGluGluAspAlaLeuAsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeu 561  
 DB 1693 ACAAGAGAAATGCTCTCATCATATCAATGATGCAATGATCAACGATATATCAAAAGATTA 1752  
 QY 562 AsnTrpGluLeuLeuIleProAspAsnSerValProIleThrSerLysLysHisAlaPhe 581  
 DB 1753 AATTGGGAATCTTCAACCCGATGACATATTTCAAGTACGTCGAGAAACATCTTAT 1812  
 QY 582 AspIleSerArgValTrpHisIleGlyTyrlArgTyrlArgAspGlyTyrlSerPheAlaAsn 601  
 DB 1813 GAGATTAACAGAGCTTTCACCACTTACAAATATAGATGGCTTACGCGTTCGCACT 1872

QY 602 ValGluThrIysSerLeuValMetArgThrValIleGluProValProLeu 618  
 Db 1873 CAAGAAAGCAAACTTGTGTGAGAGAAAGGTCTTGAACAGATGCTCTT 1923

RESULT 9  
 US-09-887-586A-29  
 / Sequence 29, Application US/09887586A  
 / Patent No. US20020094556A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Chappell, Joseph  
 / APPLICANT: No. US20020094556A1, Joseph P.  
 / APPLICANT: Scarke, Courtney M.  
 / APPLICANT: Manna, Kathleen R.  
 / TITLE OF INVENTION: SYNTASES  
 / FILE REFERENCE: 07678-025001  
 / CURRENT APPLICATION NUMBER: US/09/887,586A  
 / CURRENT FILING DATE: 2001-06-22  
 / PRIOR APPLICATION NUMBER: 09/398,395  
 / PRIOR FILING DATE: 1999-09-17  
 / PRIOR APPLICATION NUMBER: 60/130,628  
 / PRIOR FILING DATE: 1999-04-22  
 / PRIOR APPLICATION NUMBER: 60/150,262  
 / PRIOR FILING DATE: 1999-08-23  
 / NUMBER OF SEQ ID NOS: 58  
 / SOFTWARE: FastSeq for Windows Version 3.0  
 / SEQ ID NO 29  
 / LENGTH: 2196  
 / TYPE: DNA  
 / ORGANISM: Abies grandis  
 / FEATURE:  
 / NAME/KEY: CDS  
 / LOCATION: (69)...(1949)  
 / OTHER INFORMATION: myrcene synthase  
 US-09-887-586A-29

Alignment Scores:  
 Pred. No.: 6,3e-249 Length: 2196  
 Score: 2200.50 Matches: 427  
 Percent Similarity: 80.38% Conservative: 77  
 Best Local Similarity: 68.10% Mismatches: 106  
 Query Match: 67.69% Indels: 17  
 DB: 9 Gaps: 7

US-10-025-145A-65 (1-618) x US-09-887-586A-29 (1-2196)

QY 1 MetAlaIleuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15  
 Db 69 ATGGCTCTGTGTTCTATCTACCCGCTTGCGAATCTTGCTGCGCAAGTCGTGATC 128

QY 16 SerSerSerHisGluIleLysAlaLeuArgArgThrIleProThrIleGluIleCysArg 35  
 Db 129 AATTCAATTCATGAAACATTAAGCCTCCCTTATGAAACATCCCAATCTTGAAATGCGTAG 188

QY 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55  
 Db 189 CGAGGGAATCTGTGACCGCTTCCATGAGCACAAGTTGGCCACCGCTGCACTGATGAT 248

QY 56 SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTyrAspAspAspPheIleGln 75  
 Db 249 GGTGTACAAAGACGATAGTACATACCAATCTCAATATCTGAGCAGATATTTCATACAG 308

QY 76 SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95  
 Db 309 TCTCTCA---TCAACCGCTTATGAGGAACCTCTTACCGGAAACGTGTGAGGATTAAAT 365

QY 96 GlyIleValIleLysAspIleLeuPheAsnPheLysSerLeuGluAspGlyGly----- 112  
 Db 366 GTGAGGTAAAGAGATGATA---TTCAATTCATGATGATGATGATGATGATGATGATGAT 422

QY 113 -----AsnAspPheLeuGlnArgLeuLeuValAspAspValGluArgLeuGlyIle 130  
 Db 423 TCCTTTAATGATCTATGCAACGCTTGGATATGATGATGATGATGATGATGATGATGAT 482

QY 131 AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150  
 Db 483 GCTAGACATTTCAAGAACGAGATTAACATCAGCTCTGATTTATGTTTCCTGATCTGGAG 542

QY 151 GluLysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170  
 Db 543 GAAAGCGCATGTGATGTGGAGACAGACATATTGTTATCTATCTCACTCAACTCAGCGCTG 602

QY 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValIleAsnValPhe 190  
 Db 603 GCGTTTGAACCTCTTGCATTTACGCGGTACATGTTATCTCAAGAGGTTTAAAACTTTT 662

QY 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGlyGluIleArg 210  
 Db 663 CAAGATCAAAATGACAGCTTTGATGTGCTCCCGCGT---CAGACAGAGGTGATGATGAC 719

QY 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp 230  
 Db 720 AGCGTTCTTAACCTTAATACGGGCTTCCCTCATGCTCCCTGCTGAGAAAGTTATGGA 779

QY 231 GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer 250  
 Db 780 GAGCTGAATCTTCTCCACAAATATTTGAAAGAGCTTACAAAAGATTCCAGTCTCC 839

QY 251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTyrPheIleThrAsnLeuPro 270  
 Db 840 GCT---CTTTCACAGAGATAAAGTTGTTATGAGAAATATGCTGCGCACAAATTTGCCA 896

QY 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLys-----Asn 287  
 Db 897 AGATTGGAAGCAAGAAATTAACATACACATTGGAAGAACCTTACAAAAGATTCCAGTCTCC 956

QY 288 LysAsnAlaAlaGluLysLeuLeuGluLeuAlaLysLeuGluPheAsnIlePheHisSer 307  
 Db 957 AAAAATGCTGGAGAGAGCTTTTGAACCTTGCAAAATTTGAGTTCAATATTATTAATCTCC 1016

QY 308 LeuGlnGluArgGluLeuLysHisValSerArgTyrTyrLysAspSerGlySerProGlu 327  
 Db 1017 TTACACAAAGAAAGAAATTAACATATCTTTGAGATGTGTGAGAAAGTCGATTTGCCCTAA 1076

QY 328 MetThrPheCysArgHisArgHisValGlyTyrTyrAlaLeuAlaSerCysIleAlaPhe 347  
 Db 1077 TTGACATTTGCTGGGATGATGATGATGGAATTCATCACTTTGGCTCTTGTATTTGCCATT 1136

QY 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal 367  
 Db 1137 GACCCAAAACATTTGCACTTACAGACTAGGCTTCCGCAAAATGTGTATCTTGCACAGATT 1196

QY 368 LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387  
 Db 1197 TTGAGCAGATATTAGACACTTTTGAACGATGAGACGACTTGAACCTTTCATCATCTGCA 1256

QY 388 IleLysArgTyrAspProSerSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407  
 Db 1257 ATTTAAGAGTGAATTCATCAGATAGAACACTTTCAGAAATATATGAAATGTGTGTAC 1316

QY 408 MetMetValTyrHisThrValAsnGluMetAlaAspValAlaGluLysAlaGlnGlyArg 427  
 Db 1317 ATGTCGTGTTTGAACCTGTAATATGAATCAACAGACAGAGCGGAGAGAGCTCAAGGAGAA 1376

QY 428 AspThrLeuAsnTyrAlaArgGlnAlaTyrGluAlaCysPheAspSerTyrMetGlnGlu 447  
 Db 1377 AACACTCTCACTATGTTGAAAGCGCTTGGAGGCTTATTATTGATTCATATATGGAAGAA 1436

QY 448 AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGlyGluTyrTyrLeuGluAsnGlyLys 467  
 Db 1437 GCAAAATGATCTCAATAGTGTATCTGCAATGTTTGAAGAGTATCATGATGATGAGGAAA 1496

QY 468 ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe 487  
 Db 1497 GTGAGCTCTGATATCGGTGAGCAACATTGCAACCATCTTCACTTTGAATGATGAGCTT 1556



```
QY 368 LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluPheThrAlaThr 387
Db 1197 TTGGACGATATATTACGACACCTTTGGACCATTTAGACGCTTGAACCTTCACTCGCACTCGCA 1256
QY 388 TLeysArgTyrAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407
Db 1257 ATTAAAGATGAGATTCATCAGAGATAGAACACCTTCAGAAATATATAGAAATGCTGTAC 1316
QY 408 MetMetAlaTyrThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427
Db 1317 ATGGTCTGTTTGAACCTGTAATGAACTGACACGAGGCGGAGAGAACTCAAGGAGCA 1376
QY 428 AspThrLeuAsnTyrAlaArgGlnAlaTyrGluAlaCysPheAspSerTyrMetGlnGlu 447
Db 1377 AACACTCTCACTATGTTGCAAAAGCTTGGAGGCTTATTTGATTCATATATGAGAA 1436
QY 448 AlaLysTyrPheAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys 467
Db 1437 GCAAATGATCTCTATGATGTTTCTGCGCAATGTTGAAAGATACCATGAGATGGGAA 1496
QY 468 ValSerSerAlaHisArgProCysAlaLeuGlnProLLeuThrLeuAspIleProPhe 487
Db 1497 GTAGCTCTGCATATGCGGTAGCAACATGCAACCCATCTCTGATGATGATGATGATG 1556
QY 488 ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle 507
Db 1557 CCGATTACATCTTGAAGGAATGATTTTCCATCCAGGTTCAATATATTTGGCATGTC 1616
QY 508 TLeuLysLeuLysGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGlnGlu 527
Db 1617 TTCCTTGGCTACGAGGTAGACACACGCTGCTACAAAGCGCATAGGATGTTGTAAGA 1676
QY 528 AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGluGluAspAlaLeu 547
Db 1677 GCTTGTGTATATCATGTTATATGAAAGCAATCTCGATCAACCCAAAGATGCGCTTC 1736
QY 548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTyrGluLeuLys 567
Db 1737 ATCATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1796
QY 568 ProAspAsnSerValProIleThrSerLysLysHisAlaPheAspIleSerArgValTyr 587
Db 1797 TCCACACGACAAATATTCATGCTGGCCAAAGAAATGCTTTGACATACAAAGAGCTTC 1856
QY 588 HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerLeu 607
Db 1857 CACCATCTCTACATATATGAGATGCTTTAGTGTGCGCAACAGAAACAAAAATTG 1916
QY 608 ValMetArgThrValIleGlu 614
Db 1917 GTTATGAAACACCTCTTGA 1937

RESULT 11
US-09-900-797-29
; Sequence 29, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A11, Joseph P.
; APPLICANT: Starks, Courtney M.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 2196
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (69)...(1949)
; OTHER INFORMATION: myrcene synthase
US-09-900-797-29

Alignment Scores:
Pred. No.: 6,3e-249 Length: 2196
Score: 2200.50 Matches: 427
Percent Similarity: 80.38% Conservative: 77
Best Local Similarity: 68.10% Mismatches: 106
Query Match: 67.69% Indels: 17
DB: 10 Gaps: 7

US-10-025-145a-65 (1-618) x US-09-900-797-29 (1-2196)

QY 1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15
Db 69 ATGGCTCTGTTCTTCTTCTACCGCTTGGAAATCTTGCCTGGCAAGTCCTGATC 128
QY 16 SerSerSerHisGluIleLysAlaLeuArgTyrThrIleProThrLeuGlyIleCysArg 35
Db 129 AGTTCAATTCATGAAATGAAAGCTCCCTATAGAACAAATCCAAATCTTGAAATGCGTAG 188
QY 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55
Db 189 CAGGGAATATCTGTACAGCTCTTCATGACCATCATGTTTGCCACCGCTGATGATGAT 248
QY 56 SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTyrAspAspPheIleGln 75
Db 249 GGTATCAAAAGACGATAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 308
QY 76 SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95
Db 309 TCTCTA---TCAAGCCTTATGAGGAAACCTCTTACCAAGAGCTCTGAGATTAATT 365
QY 96 GlyGluValLysAspIleMetPheAsnPheLysSerLeuGluAspGlyLys----- 112
Db 366 GTGAGGTAAAGAAATGATA---TTCAATTCATGATCTGATGATGATGATGATGATGATG 422
QY 113 -----AsnAspLeuLeuGlnArgLeuLeuValAspAspValGluArgGluIle 130
Db 423 TCCCTTATATGATCTCATGACGCTTTCGATAGTCGATAGGCTTGGAGATG 482
QY 131 AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTyrAsn 150
Db 483 GCTAGACATTTCAAGACGAGATTAACATGATGATGATGATGATGATGATGATGATGATG 542
QY 151 GluLysGlyIleGlyCysGlyValArgLeuSerValAlaThrAspLeuAsnSerThrAlaLeu 170
Db 543 GAAAGCGCATTTGATGATGAGAGACAGATATGTTACTGATCTCAACTCAACTCGCTTG 602
QY 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190
Db 603 GGGTTTGCAGCTCTTCATACACCGGTACACTGATCTCCAGAGAGTTTAAAGCTTTT 662
QY 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluGlyGluIleArg 210
Db 663 CAAGATCAAAATGACACTTTGTATGCTCCCGCT---CAGACAGAGGATGATCAGA 719
QY 211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluLysValMetAsp 230
Db 720 AGCGTTTAACTTATATCGGCTTCCTCATGCTTCCCTGGGAGAAAGTTATGAGA 779
QY 231 GluAlaGluThrPheSerThrLysTyrLeuArgGluAlaLeuGlnLysIleProAlaSer 250
Db 780 GAAAGCTGAAATCTTCTCCACAAAGATATTTGAAAGAGCTCTACAAAGATTCAGCTGCC 839
```

```

QY 251 SerIleLeuSerIleuGluIleArgaspValLeuGluTyrGlyTyrPheIsthrAsnLeuPro 270
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 840 GCT---CTTTCACAGAGATAAAGTTGTATGAAATATGGCGGACACAAATTTGGCA 896

QY 271 ArgLeuGluIleArgsnTyrMetAspValPheGlyGlnIsthrLys-----Asn 287
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 897 AGATTGGAAGCAAGAAATTACATATGACACACTTGAGAGACACCGATGATGGCTCAAT 956

QY 288 LysAsnAlaIleGluIleuLeuGluLeuAlaLysLeuGlnPheAsnIlePheHisSer 307
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 957 AAAAATGCTGGAGAAAGCTTTTAGAAGCTTGCAAAATTTGAGTTCAATATATTAACTCC 1016

QY 308 LeuGlnGluArgGluLeuLysHisValSerArgTyrTrpLysAspSerGlySerProGlu 327
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1017 TTACAAACAAAGAAATTCACATATCTTTTGAAGATGGGAGAAAGATGGATTGGCTTAA 1076

QY 328 MetThrPheGlyArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1077 TTGACATTTGCTGGCGATGCGTCATGTGGAATTCACACTTTGGCTCTTGTATTTGCCAT 1136

QY 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLeuIleThrVal 367
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1137 GACCCAAACATTTGCTGATTCAGACTAGGCTTGCCAAATGTGTCTATCTTGTACAGCT 1196

QY 368 LeuAspAspMetCysAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1197 TTGGACATATTTACGACACTTTTGGAAAGATTCAGACGCTTGAACCTTCACATCTGCA 1256

QY 388 IleLysArgThrPheAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1257 ATTAAAGATGGAATTCATTCAGAGATAGAACACTTCCAGAAATATAGAAATGTGTAC 1316

QY 408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1317 ATGGTCTGTTTGAACCTGTAAATGAACTGACAGAGAGCGGAGAAAGACTCAAGGAGGA 1376

QY 428 AspThrLeuAsnTyrAlaArgGlnAlaTyrGlnAlaCysPheAspSerTyrMetGlnGlu 447
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1377 AACACTCTCAACTATGTTCGAAGGCTTGGAGGCTTATTTGATTCATATATGAAAGAA 1436

QY 448 AlaLysTrpIleAlaThrGlyTyrLeuProThrPheGlnGluTyrLeuGluAsnGlyLys 467
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1437 GCAGAAATGATCTCTAATGTTATCTGCCAATGTTTGAAGGATGCATGAGAAATGGGAAA 1496

QY 468 ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe 487
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1497 GTGAGCTCTGCATATCGCGTAGCAACATTCGCAACCCATCTCCTTGAATGCAATGGCTT 1556

QY 488 ProAspHisIleLeuLysGluValAspPheProSerLysLeuAsnAspLeuIleCysIle 507
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1557 CCTATATACATCTTGAGAGGAATGATTTCCATCCAGGTTCAATGATTTGGCATCGTCC 1616

QY 508 IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGlnGlu 527
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1617 TTCCTTGGCTACAGAGTACACACAGCGCTGTCACAAAGCCGATGAGGATCGGTGAGAAA 1676

QY 528 AlaSerSerIleSerCysTyrMetLysAspAsnProGlyLeuThrGlnGluAspAlaLeu 547
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1677 GCTTCGTATATATCATGTTATATGAAACAACTCTGATCAACCGAAGAAAGATGCCCTC 1736

QY 548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLys 567
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1737 AATCATATCAATGCCATGTCATGACATCAATCAAGAAATTAATTTGGAACTTCTAAGA 1796

QY 568 ProAspAsnSerValProIleThrSerLysHisAlaPheAspIleSerArgValTrp 587
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1797 TCACAAGCAATATTCGAATGCTGCCAAGAAACATGCTTTGACATTAACAAGACTCTC 1856

QY 588 HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrLysSerIleu 607
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1857 CACCATCTCTACATATATCGAGATGCTTATGTGTGCCAACAAGAAACAAAAAATTTG 1916

QY 608 ValMetArgThrValIleGlu 614

```

```

Db 1917 GTTATGAAACACTCCTTGAA 1937

RESULT 12
US-09-893-820-29
; Sequence 29, Application US/09893820
; Publication No. US20040053386A1
GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20040053386A11, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ. ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ. ID NO 29
; LENGTH: 2196
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (69)---(1949)
; OTHER INFORMATION: myrcene synthase
US-09-893-820-29

Alignment Scores:
Pred. No.: 6.3e-249 Length: 2196
Score: 2200.50 Matches: 427
Percent Similarity: 80.38% Conservative: 77
Best Local Similarity: 68.10% Mismatches: 106
Query Match: 67.69% Indels: 17
DB: Gaps: 7

US-10-025-145a-65 (1-618) x US-09-893-820-29 (1-2196)

QY 1 MetAlaLeuLeuSerIleThrProLeuValSerArgSerCysLeu----- 15
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 69 ATGGCTCTGTTTATCTTACCTCACCGTTGGCGAAATCTTGCTGGCGCAAGTCGTTATC 128

QY 16 SerSerSerHisGluIleLysAlaLeuArgArgThrIleProThrLeuGlyIleCysArg 35
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 129 AGTTCATTTCTGAAACATTAAGCTTCCTCCATTAAGAACATCCCAAAATCTTGAATGGGTAG 188

QY 36 ProGlyLysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 189 CGAGGAAATCTGTCAGCGCTTCCATGAGCATGAGTTGGCCAGCGCGCATGATGAT 248

QY 56 SerValGlnArgGlyValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIleGln 75
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 249 GGTGTACAAACCGATGAGTGTACCTACCATTCACATATCTGGAGAGATGATTTCTACAG 308

QY 76 SerLeuIleSerTrpProTyrGlyValaProAspTyrArgGluArgAlaAspArgLeuIle 95
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 309 TCTCTA---TCAACGCTTATGGGGAACCTCTTACAGGAACGTCGAGAGATTAAAT 365

QY 96 GlyGluValLysAspIleMetPheAsnPheLysSerIleuGluAspGlyGly----- 112
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 366 GTGAGGTAAAGAGATA---TTCATTCATTAATGTAACCTGATGATGAGAGATTAAATGAGT 422

QY 113 -----AsnAspLeuLeuGlnArgLeuLeuValaAspAspValGluArgLeuGlyIle 130
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 423 TCCTTTAATGATCTCATCTCAACGCGCTTGGATGAGTGAAGCGTTGAACGTTTGGGAGATA 482

```

QY	13	AspArgHis	Sphe	Leu	Val	Leu	Val	Leu	Val	Asp	Val	Val	Asn	Val	Trp	Phe	150
		483	GCTAG	CA	TTT	CAG	AA	CGA	GA	TAA	CA	TAC	GCT	CGA	TAT	AGT	TTT
			CGG	AG	542												
QY	151	Glu	Val	Gly	Gly	Leu	Arg	Gly	Arg	Gly	Val	Val	Thr	Asp	Leu	Val	170
		543	GA	AA	CGC	AT	TGC	AT	TGC	GA	GAC	AG	TAT	TGT	TAC	GAT	CTC
			CA	CT	CA	CT	CA	CT	CA	CT	CA	CT	CA	CT	CA	CT	CA
QY	171	Gly	Leu	Arg	Thr	Leu	Arg	Leu	His	Gly	Val	Thr	Val	Ser	Ser	Val	190
		603	GGA	TTT	CGA	ACT	CTT	CGA	TTT	CA	CGG	TAC	CGG	TAT	CTC	CAG	AGG
			TTT	AAA	GCT	TTT											662
QY	191	Lys	Asp	Lys	Asn	Gly	Gln	Phe	Ser	Ser	Trp	Ala	Asn	Leu	Gln	Gly	210
		663	CA	AG	AT	CA	AA	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	AT
			CGA	AT	CA	AA	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	AT	719
QY	211	Gly	Val	Leu	Asn	Leu	Phe	Arg	Ala	Ser	Leu	Val	Ala	Phe	Pro	Gly	230
		720	AG	CG	TT	CT	TAA	CT	TAA	CT	TAA	CT	TAA	CT	TAA	CT	779
			CGA	AT	CA	AA	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	AT	839
QY	231	Glu	Ala	Gln	Thr	Phe	Ser	Thr	Lys	Trp	Leu	Ala	Gln	Ala	Leu	Gln	250
		780	GAG	CT	GA	AT	TCT	TCT	CC	CA	CA	GA	TAT	TG	AA	AG	CT
			CA	CT	CA	CT	CA	CT	CA	CT	CA	CT	CA	CT	CA	CT	839
QY	251	Ser	Leu	Ser	Leu	Gln	Leu	Leu	Arg	Asp	Val	Leu	Gln	Val	Trp	His	270
		840	GCT	---	CTT	T	CA	CA	AG	AT	TAA	AG	TTT	GTT	TAT	G	896
			CA	AA	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	956
QY	271	Arg	Leu	Gln	Val	Ala	Arg	Asn	Trp	Met	Asp	Val	Phe	Gly	Gln	His	287
		897	AG	AT	T	GA	AT	T	GA	AT	T	GA	AT	T	GA	AT	956
			CA	AA	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	1016
QY	288	Lys	Asn	Ala	Ala	Gly	Leu	Ser	Leu	Gln	Val	Ala	Ser	Leu	Gln	Phe	307
		957	AAA	AT	T	GC	T	GA	AT	T	GA	AT	T	GA	AT	T	307
			CA	AA	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	1016
QY	308	Leu	Gln	Val	Arg	Gly	Leu	Leu	Lys	His	Val	Ser	Arg	Trp	Trp	Lys	327
		1017	TTA	CA	CA	AA	AT	T	GA	AT	T	GA	AT	T	GA	AT	1076
			CA	AA	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	1136
QY	328	Met	Trp	Phe	Cys	Arg	His	Ala	Arg	His	Val	Gln	Val	Trp	Trp	Ala	347
		1077	TTG	CA	AT	T	GC	T	GA	AT	T	GA	AT	T	GA	AT	1136
			CA	AA	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	1196
QY	348	Glu	Pro	Gln	His	Ser	Gly	Phe	Arg	Leu	Gly	Phe	Thr	Lys	Met	Ser	367
		1137	GAC	CA	AA	AT	T	GC	AT	T	GA	AT	T	GA	AT	T	367
			CA	AA	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	AT	TGC	AG	1196

QY	468	ProAspHisIleLeuValAlaAspHeProSerLysLeuAnaSpLeuIleCysIle	507
Dd	1557	CCTGATTCACATCTTGAAAGCGAATTGATTTTCCATCCAGCTCCAAATGATTTGGCATCTGCC	1616
QY	508	IleLeuArgLeuArgGlyAspThrArgCysTyrllysAlaAspArgAlaArgGlyGlutGlu	527
Dd	1617	TTCCTTCGGGTACAGAGTGACACACGCCTGTACAAGGCCGATAGGAGTCGTGGTGAAGA	1676
QY	528	AlaSerSerIleSerCysTyrlMetLysAspAnbProGlyLeuThrgLugLuAspAlaLeu	547
Dd	1677	GCTTCGTGTATATCATGTTATATATGAAGAACAATCCTGGATCAACGGAAGAAAGATGCCCTC	1736
QY	548	AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuLeuLys	567
Dd	1737	AATCATATCAATGCCATGTCATGAATGAACAATAATCAAAGATTMAATTGGGAATCTCTAGA	1796
QY	568	ProAspAsnSerValProIleThrSerLysIshAlaPheAspIleSerArgValTrp	587
Dd	1797	TCCAACGACATATTCCATCTCTGCGCCAGAAAATGCTTTTGACATPACACAGAGCTTC	1856
QY	568	HisHisGlyTYrArgTYrArgAspGlyTYriserPheAlaAsnValGluThrLysSerLeu	607
Dd	1857	CACCATCTCTCATATATACGATGGCGTTTAGTGTGTGCCAACAGAAACAAAAAATTG	1916
QY	608	ValMetArgThrValIleGln	614
Dd	1917	GTATGGAACACTCTCTTGAA	1937

```

RESULT 13
US-10-041-007-25
; Sequence 25, Application US/10041007
; Publication No. US20020164736A1
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seiichi P.T.
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase
; FILE REFERENCE: P03081US1
; CURRENT APPLICATION NUMBER: US/10/041, 007
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/259, 881
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 2196
; TYPE: DNA
; ORGANISM: Abies grandis
US-10-041-007-25

```

Alignment Scores:	
Pred. No.:	6.3e-249
Score:	2300.50
Percent Similarity:	80.3%
Best Local Similarity:	68.1%
Query Match:	67.6%
DB:	14
Gaps:	7
Length:	2196
Matches:	427
Conservative:	77
Mismatches:	106
Indels:	17

US-10-025-145A-65 (1-618) X US-10-041-007-25 (1-2196)

QY	1	MetAlaIeuLeuSerIleThrProIeuValSerAtgSerCysLeu-----	15
Db	69	ATGGCTCTGATTCATCTCACCGTGGCTTCGAATCTTCCTGCGAAGTCGTGATC	122
QY	16	SerSerSerHisGluIleValAlaLeuArgatGThrIleProThrIleGlyIleCysArg	35
Db	129	AGTCATATTCATGAACATATAGCCCTCCATATGACATCCAAATCTTGGAATGCGTAGG	188
QY	36	ProGlyIlySerSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp	55
Db	189	CGAGGAAATCTGTACACGCTTCCATGTAGCAATCACTTGGCCACCGGTGCACCTGATGAT	248
QY	56	SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuThrPaspPaspPheIleGln	75

Db 249 GGTGTACAAAGCGATAGTACTACCATTCATATCTGGACGATGATTCATACAG 308  
 Qy SerLeuIleSerThrProTyrGlyAlaProAspTyrArgGluArgAlaAspArgLeuIle 95  
 Db 309 TCTCTA---TCACGCCCTTATGGGGAAACCTCTTACACAGAACGTCGTGAGATTAAT 365  
 Qy 96 GlyGluValIlyAspIleMetPheAsnPhelSerLeuGluAspGlyGly----- 112  
 Db 366 GTGAGGTAAAGAGATA---TTCAATTCAATGACTGATGATGAGATTAATATAGT 422  
 Qy 113 -----AsnAspLeuLeuGlnArgLeuLeuValAspAspValGluArgLeuGlyIle 130  
 Db 423 TCCTTTAATGATCTGACGAAACGCTTTGGATTCGATAGCGTTGAAACGTTGGGATA 482  
 Qy 131 AspArgHisPheIlySerGlyIleIysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150  
 Db 483 GCTAGACATTTCAAGACGAGATACATCAGCTCGATTATGTTTCCGTTACTGGAGAG 542  
 Qy 151 GluIlySerGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170  
 Db 543 GAAACGCGCATGGATGGAGAGACAGATATGTTGTTGATCTCACTCACTCACTGCGCTG 602  
 Qy 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValIleuAsnValPhe 190  
 Db 603 GGGTTTGAACCTCTTCGATTAACGCGGTACACTGTATCTCCAGAGGTTTAAAGCTTTT 662  
 Qy 191 LysAspIlyAsnGlyIlyPheSerSerThrAlaAsnIleGlnIleGluGlyGluIleArg 210  
 Db 663 CAAGATCAAAATGACAGTTTGTATGCTCCGCCGCT---CAGACAGAGGGGAGATACAGA 719  
 Qy 211 GlyValLeuLeuLeuPheArgAlaSerLeuValAlaPheProGlyGluIlyValMetAsp 230  
 Db 720 AGCCTTCTTAATTAATCGGGCTTCCCTCATTCGCTTCCCTGATGAGAAAGTTATGAA 779  
 Qy 231 GluAlaGluThrPheSerThrIleTyrLeuArgGluAlaLeuGlnIlyIleProAlaSer 250  
 Db 780 GAACCTGAAATCTCTCCACAAAGATATTGAAGAAGCTTACAAAGATTTCCAGTCTCC 839  
 Qy 251 SerIleLeuSerLeuGluIleArgAspValLeuGluTyrGlyTyrPheIsthAsnLeuPro 270  
 Db 840 GCT---CTTTCACAGAGATTAAGATTGTTGTTATGGAATGCTGCGACACAAATTTGGCA 896  
 Qy 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisIsthIlyS-----Asn 287  
 Db 897 AGATTGGAGCAAGAAATTAATACATACACACTTGAGAAAGACACCAAGCATGCTCATAT 956  
 Qy 288 LysAsnAlaAlaGluIlyLeuLeuGluLeuAlaIlySLeuGluPheAsnIlePheHisSer 307  
 Db 957 AAAAATGCTGGGAAGAAGCTTTAGAACTTGCAAAATTCGAGTTCAATATTTAACTCC 1016  
 Qy 308 LeuGlnGluArgGluLeuLeuHisIsthValSerArgTyrTrpIlyAspSerGlySerProGlu 327  
 Db 1017 TTACAAACAAAGGAATTAATATCTTTTGAGANTGTGGAAGAAAGATCGAATTTGCTTAA 1076  
 Qy 328 MetThrPheCysArgHisArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347  
 Db 1077 TTGACATTTGCTCGGACATCGTCATGTGAAATTTCTACACTTGGCCTCTTGATTTGCCATT 1136  
 Qy 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrIlySMetSerHisLeuIleThrVal 367  
 Db 1137 GACCCAAACATTTGCACTTCGACTAGCTGCGCAAAATGTGTCATCTTGTCACAGTT 1196  
 Qy 368 LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluPheThrAlaThr 387  
 Db 1197 TTGACGATATTTAGACACTTTTGAACGATTCGACGACTTGAACTCTTCACATCTGCA 1256  
 Qy 388 IleIlyAspGlyTrpAspProSerAlaMetGluCysLeuProGluTyrMetIlySerGlyValTyr 407  
 Db 1257 ATTAAAGATGGAATTCATCAGAGATGAACACCTTCCAGATATATGAAAGTGTGTAAC 1316  
 Qy 408 MetMetValTyrHisIsthValAsnGluMetAlaArgValAlaGluIlyAlaGlnIlyArg 427  
 Db 1317 ATGTCGTGTTGAAACTGTAAATGAACCTGACACAGAGAGCGGAGAAAGCTCAAGGAGAGA 1376

Qy 428 AspThrLeuAsnTyrAlaArgGluAlaTrpGluAlaCysPheAspSerTyrMetGlnGlu 447  
 Db 1377 AACACTCTCAACTAGTGTCCAAAGGCTTGGAGGCTTATTTGATTAATATGGAAGAA 1436  
 Qy 448 AlalysTrpIleAlaThrGlyTyrLeuProThrPheGluGluTyrLeuGluAsnGlyLys 467  
 Db 1437 GCAAAATGATCTCTAATGTTATCTTCGCAATGTTTAAAGAGTACATGAAATGGGAAA 1496  
 Qy 468 ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe 487  
 Db 1497 GTGAGCTCTGCATATCGCGTAGAACCAATTCACACCCATCTCACTTTGAAATGCATGCTT 1556  
 Qy 488 ProAspHisIleLeuIlySValAlaAspPheProSerIlySLeuAsnAspLeuIleCysIle 507  
 Db 1557 CTTGATTAATCTTTGAAGGAATGATTTTCCATCCAGGTTCAATGATTTGGCATGCTCC 1616  
 Qy 508 IleLeuArgLeuArgGlyAspThrArgCysTyrIlyAlaAspArgAlaArgGlyGluGlu 527  
 Db 1617 TTCCTTCGCTACGAGGCTGACACAGCGCTGACAGAGCCGATAGGAGATCGGTGGAAGAA 1676  
 Qy 528 AlaserSerIleSerCysTyrMetIlyAspAspProGlyLeuThrGluAspAlaLeu 547  
 Db 1677 GCTTCGTATATCATGTTATATGAAGACATCTGATCAACCGAAGAGATGCCCTC 1736  
 Qy 548 AsnHisIleAsnPhenMetIleArgAspAlaIleArgGluLeuAsnTrpGluLeuIlyS 567  
 Db 1737 AATCATTCATATGCGCATGCTCAATGACATATCAAGATTAATTTGGGAACCTCTAAGA 1796  
 Qy 568 ProAspAsnSerValProIleThrSerIlySHisAlaPheAspIleSerArgValTrp 587  
 Db 1797 TCCAAAGCAATATTTCCAAATGCTGCGCAAGAAACATGCTTTTGCAATACAAAGGCTCTC 1856  
 Qy 588 HisIsthGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGluThrIlySerLeu 607  
 Db 1857 CACCATCTCAATATATGAGATGCGCTTTATGTTGCCACAAAGAAACAAAAATTG 1916  
 Qy 608 ValMetArgThrValIleGlu 614  
 Db 1917 GTTATGAAACACTCCTTGAA 1937

RESULT 14  
 US-10-025-145A-1  
 : Sequence 1, Application US/10025145A  
 : Publication No. US20030175861A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Croteau, Rodney B.  
 : APPLICANT: Bohnmann, Joerg  
 : APPLICANT: Steele, Christopher L.  
 : APPLICANT: Phillips, Michael A.  
 : TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)  
 : FILE REFERENCE: WSUR118414  
 : CURRENT APPLICATION NUMBER: US/10/025,145A  
 : CURRENT FILING DATE: 2002-06-28  
 : PRIOR APPLICATION NUMBER: US 09/360,545  
 : PRIOR FILING DATE: 1999-07-26  
 : PRIOR APPLICATION NUMBER: PCT/US98/14528  
 : PRIOR FILING DATE: 1998-07-10  
 : PRIOR APPLICATION NUMBER: US 60/052,249  
 : PRIOR FILING DATE: 1997-07-11  
 : NUMBER OF SEQ ID NOS: 107  
 : SOFTWARE: PatentIn version 3.1  
 : SEQ ID NO 1  
 : LENGTH: 2196  
 : TYPE: DNA  
 : ORGANISM: Abies Grandis  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: (69)..(1952)  
 : OTHER INFORMATION:  
 : US-10-025-145A-1

Alignment Scores:

Pred. No.: 6.3e-249 Length: 2196  
 Score: 2200.50 Matches: 427  
 Percent Similarity: 80.38% Conservative: 77  
 Best Local Similarity: 68.10% Mismatches: 106  
 Query Match: 67.69% Indels: 17  
 DB: 15 Gaps: 7

US-10-025-145A-65 (1-618) x US-10-025-145A-1 (1-2196)

QY 1 MetAlaLeuSerLleuHrProLeuValSerArgSerCysLeu----- 15  
 Db ATGGCTGCTGTTTCTATCTCACCCTGGCTTGAATCTTGTCTGGCCAGAGTGGTATC 128  
 QY 16 SerSerSerHsGluIleLysAlaLeuArgHrHrLleProHrLleuGlyLeCysArg 35  
 Db ATTTCAATTATGACAACTAAGCCTCCCTATGAACAAATCCCAAACTTGGAAATGCTAGG 188  
 QY 36 ProGlyLysSerValAlaHisSerLleAsnMetCysLeuHrSerValAlaSerThrAsp 55  
 Db CGAGGGAATCTGACAGCCCTTCATGAGCATGCTTGGCCACCGCTGCACCTGATGAT 248  
 QY 56 SerValGlnArgArgValGlyAsnTyrHisSerAsnLeuTrpAspAspPheIleGln 75  
 Db GGTGTCAAAAGAGCATAGGTGACTACCATTCATATCTGGAGCGATGATTTCTATCAG 308  
 QY 76 SerLeuLleSerThrProTyrGlnGlyAlaProAspTyrArgGlnArgAlaAspArgLeuIle 95  
 Db TCTCTA---TCAAGCCTTAAGGGGAACCTCTTACAGAGACGCTGTGAGAGATTAAAT 365  
 QY 96 GlyGluValLysAspLleMetPheAsnPhelLysSerLeuGluAspGlyGly----- 112  
 Db GTGAGAGTAAAGAAGATA---TTCATTCATGACTGATGATGATGAAGATTAATAGACT 422  
 QY 113 -----AsnAspLeuGlnGlnArgLeuLeuValAspAspValGlnArgLeuGlyIle 130  
 Db TCTTAAATATCTCTCATGCAACCGCTTGGATGTCATAGCCTTGAACGTTGGGGATA 482  
 QY 131 AspArgHisPheLysLysGluIleLysThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150  
 Db GCTAGACATTTCAAGAACGATTAACATCACTGCTGATTAATGTTTCCGTTACGGGAG 542  
 QY 151 GlnLysGlyLleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170  
 Db GAATAAGGACATGATGATGTGGAGAGACATATTCTTATCTCAACTCACTGCGCTTG 602  
 QY 171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnValPhe 190  
 Db GGGTTCGAACCTCTCGATTACACGGGTACACTGTAATCTCCAGAGGTTTAAAGCTTTT 662  
 QY 191 LysAspLysAsnGlyGlnPheSerSerThrAlaAsnLleGlnLleGluGluIleArg 210  
 Db CAAGATCAAAATGACAGCTTGTATGCTCCCGGCT---CAGACAGAGGGGTGAGATCAGA 719  
 QY 211 GlyValLeuAsnLeuPheArgAlaSerLleValAlaPheProGlyGluLysValMetAsp 230  
 Db AGGCTTCTTAACCTTAATATGGGCTTCCCTCATCTGCTCCCTGTGAGAAAGTTATGAA 779  
 QY 231 GlnAlaGlnThrPheSerThrLysTyrLeuArgGlnAlaLeuGlnLysIleProAlaSer 250  
 Db GAAGCTGAATCTCTCCACAGATATTTGAAGAAGAGCTCTCAAAAGATTCCAGCTCC 839  
 QY 251 SerLleLeuSerLeuGluIleArgAspValLeuGluTyrGlyThrHisThrAsnLeuPro 270  
 Db GCT---CTTTCACAGAGATTAAGTTTGTATGAGATATGGCTGGACACAAATTTGCCA 896  
 QY 271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnHisThrLys-----Asn 287  
 Db AGATTGGAAGCAAGAAATTACATACACACTTGAAGAAGACACCGAGTGATGGCTCAAT 956  
 QY 288 LysAsnAlaAlaGluLysLeuLeuGluLeuAlaLysLeuGluPheAsnLlePheHisSer 307  
 Db AAAAATGCTGGGAAGAGCTTTTAAAGCTTGAACAAATGGAGTTCAATATATTAACTCC 1016

QY 308 LeuGlnGluArgGluLeuLysHisValSerArgTrpTrpLysAspSerGlySerProGlu 327  
 Db TTAACAACAAAGAAATTACAAATATCTTTTGAAGATGGTGAAAGAGTGGATTTTGCTTAA 1076  
 QY 328 MetThrPheCysArgHisValGlnTyrTyrAlaLeuAlaSerCysLleAlaPhe 347  
 Db TTGACATTTCTCGGACATGCTCATGAGATTTCAACCTTGGCTCTTGTATGGCAATT 1136  
 QY 348 GluProGlnHisSerGlyPheArgLeuGlyPheThrLysMetSerHisLleuHrVal 367  
 Db GACCAAAACATCTGCTGATTCAGACTAGGCTTCCCAAAATGTCTCATCTTGTCAAGCTT 1196  
 QY 368 LeuAspAspMetCysArgAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387  
 Db TTGACGATATTTAACAACACTTTTGGAAACGATGACAGCTTGAACCTTCTCATATGCA 1256  
 QY 388 IleLysArgTrpAspProSerAlaMetGluCysLeuProGluTyrMetLysGlyValTyr 407  
 Db ATTAAGAGATGAATTCATCAGATGAAACACCTTCCAGAAATTAATGAATGTGTATC 1316  
 QY 408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluLysAlaGlnGlyArg 427  
 Db ATGGTGTGTTTGAACCTGAATGAATGACATGACACAGAGCGGAGAAAGACTCAAGAGAGA 1376  
 QY 428 AspThrLeuAsnTyrAlaArgGlnAlaTyrGluAlaCysPheAspSerTyrMetGlnGlu 447  
 Db AACACTCTCAACTATGTTCCAAAGAGCTTGGAGGCTTATTTGATCATATATGAGAGA 1436  
 QY 448 AlaLysTrpLleAlaThrGlyTyrLeuProThrPheGlnGluTyrLeuGluLysGlyLys 467  
 Db GCAAAATGATCTCTAATGTTTATCTGCAATGTTTGAAGATGACATGAGAAATGGGAAA 1496  
 QY 468 ValSerSerAlaHisArgProCysAlaLeuGlnProLleuThrLeuAspLleProPhe 487  
 Db GTAGCTCTCATATGATGCGTAGACAACTTGCACACCTCTCACTTGTGATGATGCTT 1556  
 QY 488 ProAspHisLleLeuLysGluValAspPheProSerLysLeuAsnLeuLleCysIle 507  
 Db CCGATTTACATCTTGAAGAGAAATTTTCCATCCAGCTTCAATGATTTGGCATGCTCC 1616  
 QY 508 IleLeuArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAlaArgGlyGluGlu 527  
 Db TTTCTTGGCTTACGAGGTGACACACCTGCTTACAAAGCGCATAGGGATCGTGTGAAGAA 1676  
 QY 528 AlaSerSerLleSerCysTyrMetLysAspAsnProGlyLeuThrGlnGluAspAlaLeu 547  
 Db GCTTCGTATATCATGATTATATGAAGAACAACTCGGATCAACCGAAGAAATGCCCTC 1736  
 QY 548 AsnHisLleAsnPheMetLleArgAspAlaAlaArgGluLeuAsnTrpGluLeuLys 567  
 Db ATTCATATCATATGCCATGCTCATATGACATTAACAAAGAAATTAATGGGAATTTTAAAGA 1796  
 QY 568 ProAspAsnSerValProLleThrSerLysLysHisAlaPheAspLleSerArgValTrp 587  
 Db TCCAAAGACAAATTTCAATGCTGCGCAAGAAACATGCTTTTGACATTAACAAAGCTCTC 1856  
 QY 588 HisHisGlyTyrArgTyrArgAspGlyTyrSerPheAlaAsnValGlnThrLysSerLeu 607  
 Db CACCATCTCTACATATATGAGATGCTTATAGTGTGGCCACAAAGAAACAAAAATTTG 1916  
 QY 608 ValMetArgThrValLleGlu 614  
 Db GTTATGAAACACTCTTGA 1937

# RESULT 15

US-10-025-145A-31  
 ; Sequence 31, Application US/10025145A  
 ; Publication No. US20030175861A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Croteau, Rodney B.  
 ; APPLICANT: Bohlmann, Joerg  
 ; APPLICANT: Steele, Christopher L.  
 ; APPLICANT: Phillips, Michael A.

TITLE OF INVENTION: Monoterpene Synthases from Grand Fir (Abies Grandis)  
FILE REFERENCE: WSUR118414  
CURRENT APPLICATION NUMBER: US/10/025,145A  
CURRENT FILING DATE: 2002-06-28  
PRIORITY APPLICATION NUMBER: US 09/360,545  
PRIORITY FILING DATE: 1999-07-26  
PRIORITY APPLICATION NUMBER: PCT/US98/14528  
PRIORITY FILING DATE: 1998-07-10  
PRIORITY APPLICATION NUMBER: US 60/052,249  
PRIORITY FILING DATE: 1997-07-11  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 31  
LENGTH: 2205  
TYPE: DNA  
ORGANISM: Abies Grandis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (57)..(1943)  
OTHER INFORMATION:  
US-10-025-145A-31

## Alignment Scores:

Pred. No.:	1.44e-248	Length:	2205
Score:	2197.50	Matches:	427
Percent Similarity:	80.38%	Conservative:	77
Best Local Similarity:	68.10%	Mismatches:	106
Query Match:	67.59%	Indels:	17
DB:	15	Gaps:	7

US-10-025-145A-65 (1-618) x US-10-025-145A-31 (1-2205)

1 MetAlaLeuLeuSerThrProLeuValSerArgSerCysLeu----- 15  
57 ATGCTCTGGTTTCTATCTCAACCGTTGGCTTGGAAATCTTGGCGCAAGTCGTGATC 116  
16 SerSerSerHisGluIleValAlaLeuArgArgThrIleProThrLeuGlyIleCysArg 35  
117 AGTTCATTTCAAGACATTAAGCTCCCTTAAGAAATCCCAATCTTGGAAATGCGTAGG 176  
36 ProGlyIysSerValAlaHisSerIleAsnMetCysLeuThrSerValAlaSerThrAsp 55  
177 CGAGGGAATCTGTACGCCCTTCCATGACATCAAGTTTGGCCACCGTCGACCTGATGAT 236  
56 SerValGlnArgValGlyAsnTyrHisSerAsnLeuTTPAspAspPheIleGln 75  
237 GGTGTACAAAGACGATGAGTGAAGTCAATTCATTCGAGAGATGATTTCTACAG 296  
76 SerLeuIleSerThrProTyrGlyValProAspTyrArgGluArgAlaAspArgLeuIle 95  
297 TCTCTA---TCAACCATTTAGGGGAACCTTTTCCAGAAACGTCGAGAGATTAAAT 353  
96 GlyIleValIysAspIleMetPheAsnPheIysSerIleGluAspGlyGly----- 112  
354 GTGAGGATTAAGAAAGATA---TTCAATTCATGACCTGATGATGAGAAATTAATGAGT 410  
113 -----AsnAspLeuLeuGlnArgLeuLeuValAspAspValGlnArgLeuGlyIle 130  
411 TCTTTAATGATCTCATCAACGCTTTGGATAGTCGATAGCGTTGAACGTTTGGAGATA 470  
131 AspArgHisPheIysGluIleValSerThrAlaLeuAspTyrValAsnSerTyrTrpAsn 150  
471 GCTACACATTTTCAACAGAGATACATCAGCTCTGATTAATGTTTCCGTTACCGGAG 530  
151 GluIysGlyIleGlyCysGlyArgGluSerValValThrAspLeuAsnSerThrAlaLeu 170  
531 GAAACCGCATTTGATGTGGAGAGACAGTATGTACTGATCTCAACTCAACTCGCTTG 590  
171 GlyLeuArgThrLeuArgLeuHisGlyTyrThrValSerSerAspValLeuAsnAlaPhe 190  
591 GGGTTTCGAACCTTCTGATTAACGAGTACCTGTATCTCCAGAGCTTTTAAACCTTTT 650  
191 LysAspIysAsnGlyGlnPheSerSerThrAlaAsnIleGlnIleGluIleArg 210

651 CAAATCAAAATGACAGCTTTGATGCTCCCGCGT---CAGACAGAGGGTGAATCAGA 707  
211 GlyValLeuAsnLeuPheArgAlaSerLeuValAlaPheProGlyGluIysValMetAsp 230  
708 AGCCTTCTTAATCTTAATCGGGCTTCCCTCATTCCTTCCCTGGTGAAGATTAAGAA 767  
231 GluAlaGluThrPheSerThrIysTyrLeuArgGluAlaLeuGlnIysIleProAlaSer 250  
768 GAACTCAAAATCTTCTCAACAGATATTGAAAGAAAGCTCTACAAAAGATTCAGTCTCC 827  
251 SerIleLeuSerLeuGluIleArgAspValLeuGluIlyrGlyTyrPheIsthAsnLeuPro 270  
828 GCT---CTTTCACAAAGATTAAGTTTGTATGAAATAGGCTGGACACAAATTTGCCA 884  
271 ArgLeuGluAlaArgAsnTyrMetAspValPheGlyGlnIsthIys-----Asn 287  
885 AGATTGAGAGCAAAATTCATGACACACACTTGAGAAAGACACCAAGTCAGTCTCAAT 944  
288 LysAsnAlaAlaGluIysLeuLeuGluIleValIysLeuGluIysPheAsnIlePheHisSer 307  
945 AAAATGCTGGAGAAAGACCTTTTGAAGCTTCCAAATTTGAGTTCAATATTTAACTCC 1004  
308 LeuGlnGluArgGluLeuIysHisValSerArgTyrTrpIysAspSerGlySerProGlu 327  
1005 TTACACAAAAGAAATTACAAATATCTTTGAGATGTTGGAAGAGTGCATTTGGCTAAA 1064  
328 MetThrPheCysArgHisValGluTyrTyrAlaLeuAlaSerCysIleAlaPhe 347  
1065 TTGACATTTGCTCGGACGTCGATGGAATTCATCACTTTGGCTCTGTATTTGCCATT 1124  
348 GluProGlnHisSerGlyPheArgLeuGlyPheThrIysMetSerHisLeuIleThrVal 367  
1125 GACCCAAAACATTTCTGATTTAGACTAGAGCTTCCGCAAAATGTGTCACTTGTACAGTT 1184  
368 LeuAspAspMetTyrAspValPheGlyThrValAspGluLeuGluLeuPheThrAlaThr 387  
1185 TTGGACGATTTTTCGACACTTTTGAACGATTCAGAGCTTGAACCTTTCACATCTTCA 1244  
388 IleIysArgTTPAspProSerAlaMetGluCysLeuProGluIyMetIysGlyValIyr 407  
1245 ATTAGAAGATGAATTCATTCAGAGATGAAACACCTTCCAGAAATATGAAATGTGTGAC 1304  
408 MetMetValTyrHisThrValAsnGluMetAlaArgValAlaGluIysAlaGlnIyArg 427  
1305 ATGTCGCGTTTGAACCTGTAATGAATGACTGACAGAGAGGGAGAAACCTCAAGGAGA 1364  
428 AspThrLeuAsnTyrAlaArgGlnAlaTyrGluAlaCysPheAspSerTyrMetGlnGlu 447  
1365 AACACTTCACATATGTTTGAAGAGCTTGTATTTTGAATTCATATATGAGAGAA 1424  
448 AlaIysTyrPheIleAlaThrGlyTyrLeuProThrPheGluGluIyTyrLeuGluAsnGlyIys 467  
1425 GCAAAATGATCTCTAATGTATCTGCCAAGCTTGAAGATTCACAGAAATGGAGAA 1484  
468 ValSerSerAlaHisArgProCysAlaLeuGlnProIleLeuThrLeuAspIleProPhe 487  
1485 GTGAGCTCTGATATCGGCTGAGCAACATTCGAACCCATCTTCATTCATGATCATGGCTT 1544  
488 ProAspHisIleLeuIysGluValAspPheProSerIysLeuAsnAspLeuIleCysIle 507  
1545 CCTGATTCATCTTGAAGGAAATGATTTTTCATCCAGATTCAAAGATTTTGGCATCGTCC 1604  
508 IleLeuArgLeuArgGlyAspThrArgCysTyrIysAlaAspArgAlaArgGlyGluGlu 527  
1605 TTCTTTCGCTACGAGTGAGACACGCTGCTACAAAGCCGATGAGGATCGTGTAAAGA 1664  
528 AlaSerSerIleSerCysTyrMetIysAspAsnProGlyLeuThrGluGluAspAlaLeu 547  
1665 GCTTGATATATCATGTTATATGAAGACATCTCGATTCACCGAAGAAATGATCCCTTC 1724  
548 AsnHisIleAsnPheMetIleArgAspAlaIleArgGluLeuAsnTyrGluLeuLeuIys 567



**This Page Blank (uspto)**